



The  
Patent  
Office

US 99/1980

PCT/US 99 1980  
09/719379  
INVESTOR IN PEOPLE

The Patent Office  
Concept House  
Cardiff Road  
Newport  
South Wales  
NP9 1RH

REC'D 16 JUL 1999

WIPO PCT

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed

Dated

30 April 1999

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

SECRET

THIS PAGE BLANK (USPTO)

For Official use

11 JUN 1998

12JUN98 E367545-1 D02029  
PO1/7700 25.00 - 9812613.9

11 JUN 1998

Your Reference:

B45145

9812613.9

**Notes**

Please type, or write in dark ink using CAPITAL letters. A prescribed fee is payable for a request for grant of a patent. For details, please contact the Patent Office (telephone 071-483 4700).

Rule 16 of the Patents Rules 1990 is the main rule governing the completion and filing of this form.

② Do not give trading styles, for example, 'Trading as XYZ company', nationality or former names, for example, 'formerly (known as) ABC Ltd' as these are not required.

**Warning**  
After an application for a Patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977 and will inform the applicant if such prohibition or restriction is necessary. Applicants resident in the United Kingdom are also reminded that under Section 23, applications may not be filed abroad without written permission unless an application has been filed not less than 6 weeks previously in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction revoked.

**The Patent Office** Request for grant of a Patent Form 1/77 Patents Act 1977

**① Title of invention**

1 Please give the title of the invention      Vaccine

**② Applicant's details**

First or only applicant

2a If you are applying as a corporate body please give:  
Corporate Name      SmithKline Beecham Biologicals S.A.

Country (and State of incorporation, if appropriate)      Belgium

2b If you are applying as an individual or one of a partnership please give in full:

Surname

Forenames

2c In all cases, please give the following details:

Address: Rue de l'Institut 89, B1330 Rixensart, Brussels

UK postcode  
(if applicable)

Country      Belgium  
ADP number      5800974002  
(if known)

6570956 001

2d, 2e and 2f: If there are further applicants please provide details on a separate sheet of paper

<input type="checkbox"/>	<b>Second applicant (if any)</b> If you are applying as a corporate body please give: Corporate Name Ohio State Research Foundation 1960 Kenny Road, Columbus, OH 43210-1063.
Country (and State of Incorporation, if appropriate) USA 7454770001	
2e	If you are applying as an individual or one of a partnership please give in full:  Surname:  Forenames:
2f	In all cases, please give the following details:  Address:  UK postcode (if applicable) Country ADP number (if known)
<b>③</b> Address for service details  3a Have you appointed an agent to deal with your application? Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> go to 3b  please give details below  Agent's name Marcus J W Dalton  Agent's address SmithKline Beecham Corporate Intellectual Property New Horizons Court Great West Road Brentford, Middlesex  Postcode TW8 9EP Agent's ADP number 412233000	

3b: If you have appointed an agent, all correspondence concerning your application will be sent to the agent's United Kingdom address.

If you have not appointed an agent please give a name and address in the United Kingdom to which all correspondence will be sent:

Name

Address

Postcode  
number  
ADP number  
(if known)

Daytime telephone  
(if available)

**THIS PAGE BLANK (USPTO)**

7

- any applicant is not an inventor
- there is an inventor who is not an applicant, or
- any applicant is a corporate body.

8 Please supply duplicates of claim(s), abstract, description and drawings).

## 7 Inventorship

7. Are you (the applicant or applicants) the sole inventor or the joint inventors?

Please mark correct box

Yes  No  A Statement of Inventorship on Patents form 7/77 will need to be filed (see Rule 15).J.

## 8 Checklist

8a Please fill in the number of sheets for each of the following types of document contained in this application

Continuation sheets for this Patents Form 1/77

—

 Claim(s)  Description   
Abstract  Drawing(s)  ~~11~~

8b Which of the following documents also accompanies the application?

Priority documents (please state how many)

—

Translation(s) of Priority documents (please state how many)

—

Patents Form 7/77 - Statement of Inventorship and Right to Grant

—

Patents Form 9/77 - Preliminary Examination Report

—

Patents Form 10/77 - Request for Substantive Examination

—

Please mark correct box(es)

9 You or your appointed agent (see Rule 90 of the Patents Rules 1990) must sign this request.

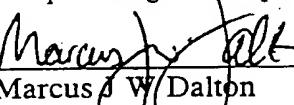
Please sign here 

A completed fee sheet should preferably accompany the fee.

## 9 Request

I/We request the grant of a patent on the basis of this application.

Signed

 Date: 11/06/98

Marcus J W Dalton

(day)

month

year

Chartered Patent Attorney  
Attorney for the Applicant

Please return the completed form, attachments and duplicates where requested, together with the prescribed fee to either;

The Comptroller or  
The Patent Office  
Cardiff Road  
Newport  
Gwent  
NP9 1RH

The Comptroller  
The Patent Office  
25 Southampton Buildings  
London  
WC2A 1AY

④ Reference number

4. Agent's or  
applicant's reference MJL/B45145  
number (if applicable)

Please mark correct box

⑤ Claiming an earlier application date

5. Are you claiming that this application be treated as having been filed on the date of filing of an earlier application?

Yes

No  go to 6



please give details below

number of earlier  
application or patent  
number

filing date (day) month year

and the Section of the Patents Act 1977 under which you are  
claiming:

15(4) (Divisional)  8(3)  12(6)  37(4)

⑥ Declaration of priority

6. If you are declaring priority from previous application(s), please give:

Country of Filing	Priority application number (if known)	Filing Date (day, month, year)

Please give the date in all number  
format, for example, 31/05/90 for  
31 May 1990

## VACCINE

## FIELD OF INVENTION

This invention relates to newly identified peptides and polynucleotides encoding 5 these peptides, and to chimeric proteins that carry these peptides. The invention also relates to a method of isolating the peptides or chimeric proteins and a vaccine composition for use in the treatment of *Haemophilus influenzae* infection.

## BACKGROUND OF THE INVENTION

10 *Haemophilus influenzae* (Hi) is a gram-negative coccobacillus and a strict human commensal. Strains of Hi are either encapsulated in a polysaccharide capsule or are non-encapsulated and are accordingly classified into typeable (encapsulated) and non-typeable (non-encapsulated) strains.

15 Encapsulated pathogenic strains of Hi cause mainly, but not exclusively, invasive disease in children under six years of age. *Haemophilus influenzae* type b (Hib), for example, is a major cause of meningitis and other invasive infections in children. Effective vaccines exist against Hib infections, and are based on producing antibodies to the polysaccharide capsule, and are therefore ineffective against non-typeable *Haemophilus influenzae* (ntHi).

20 Non-typeable *Haemophilus influenzae* (ntHi) represents the majority of the colonising strains and, although rarely invasive, are responsible for a significant proportion of mucosal disease including *otitis media*, sinusitis, chronic conjunctivitis and chronic or exacerbation of lower respiratory tract infections. Currently, approximately 30%, and as much as 62% of ntHi are resistant to penicillins. Carriage is estimated at 44 25 % in children and approximately 5 % in adults, and can persist for months. Neither the pathogenic mechanisms nor the host immunological response has been fully defined for *otitis media* caused by ntHi.

*Otitis media* is a common disease in children less than 2 years of age. It is defined by the presence of fluid in the middle ear accompanied by a sign of acute local or

systemic illness. Acute signs include ear pain, ear drainage, hearing loss whereas systemic signs include fever, lethargy, irritability, anorexia, vomiting or diarrhoea. *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* (ntHi) are the most predominant bacteria that cause the condition, accounting for 25-50%, and 15-30% of the species cultured, respectively. In addition, ntHi is responsible for 53 % of recurrent *otitis media*. Approximately 60% and 80% of children have at least one episode of the disease by 1 and 3 years of age respectively (the peak being around 10 months).

There is evidence that protective immunity does exist for ntHi, however antigenic drift in the epitopes naturally involved (outer-membrane proteins P2, P4, P6) plays a major role in the ability of ntHi to evade the immune defence of the host.

There is therefore a need for additional effective vaccines against *Haemophilus influenzae*, and particularly for vaccines against non-typeable *Haemophilus influenzae* which is not affected by the currently available Hi polysaccharide vaccines.

Fimbriae, which are surface appendages found on ntHi, are produced in 100% of the bacteria recovered from the middle ears and nasopharyngeal region of children with chronic *otitis media*. A vaccine comprised of fimbriae, a filamentous protein derived from the fimbriae of ntHi has been reported (WO 94/26304) and is useful in studying, preventing, or reducing the severity of *otitis media*. Fimbriae is homologous to the P5 outer membrane protein of ntHi that has been the subject of another patent application (EP 680765). The fimbriae P5-like protein is capable of eliciting antibodies that react to the bacteria's surface and are bactericidal (WO 94/26304). The protein has been purified and has been shown to induce an immune response against different strains of ntHi.

Existing methodologies to isolate fimbriae protein from the bacterial outer membrane were tedious and time-consuming. A strategy with other bacterial species has been to use an alternative immunogen having relatively short linear peptides. However, such alternative immunogens are of limited value since antibodies to such alternative immunogens frequently fail to recognise the native pathogen.

LB1(f) is a 19 amino-acid peptide derived from the sequence of P5-like fimbriae protein from strain ntHi1128 (occupying the region Arg117 to Gly135). This peptide was

defined initially as being a potential B cell epitope, by analysis of the primary sequence of P5-like fimbrial protein. Immunising animals with chimeric fimbrial peptides (called LB1 peptides), comprising the LB1(f) peptide, a linker peptide and a T cell epitope, induces an immune response to the P5-like fimbrial protein and reduces the concentration 5 of ntHi in animals subsequently exposed to ntHi. The LB1 chimeric fimbrial peptide is immunogenic *in vivo* and antisera generated against it was immunoreactive against both denatured and native fimbriae. The peptide was thus able to act as an effective immunogen in that it was able to generate antibodies which recognised and bound to the epitope in its native structure. This is due in part to the synthetic LB1(f) peptide 10 mimicking the coiled-coil secondary structure of the peptide within the fimbrial protein.

This invention relates to the more effective use of the LB1(f) peptide as a vaccine against *Haemophilus influenzae* strains that express the P5-like fimbrial protein (or naturally occurring variants of the protein).

## 15 SUMMARY OF THE INVENTION

It is an object of the present invention to provide groups of newly identified 20 antigenic P5-like fimbrial subunit peptides (LB1(f) peptides) of P5-like fimbrial proteins from various ntHi strains. It is a further object to provide chimeric polypeptides that carry these peptides and which induce an immunogenic response in animals to ntHi, and polynucleotides encoding such peptides and polypeptides. The invention also relates to a method of isolating the peptides or chimeric polypeptides, to a method of detecting the presence of the peptides in biological samples, and to a vaccine composition for use in the treatment of *Haemophilus influenzae* infection.

The groups of LB1(f) peptides contain peptides from about 13 to about 22 amino 25 acids in length. The peptides fall into 3 main groups (one of which contains 2 subgroups). The chimeric protein comprises one or more of the LB1(f) peptide units covalently linked to a carrier protein that additionally acts as a T-cell epitope. Preferably the carrier protein is from *Haemophilus influenzae* so it may induce an immunogenic response in animals to *Haemophilus influenzae* itself.

The invention may be more fully understood by reference to the following drawings and detailed description.

#### BRIEF DESCRIPTION OF THE DRAWINGS

5

**Figure 1:** Plasmid pMG1MCS. The DNA sequence of the multiple cloning site is given.

**Figure 2:** Plasmid pRIT14588.

**Figure 3:** Plasmid LPD-LB1-A.

10 **Figure 4:** Plasmid LPD-LB1-II. The DNA and amino acid sequences of the Group 1 (LB1-GR1) and Group 2 (LB1-GR2) LB1(f) peptides are indicated with arrows. The arrows encompass the LB1(f) within the sequence of its natural context within the p5-like fimbrin protein.

15 **Figure 5:** Plasmid LPD-LB1-III. The DNA and amino acid sequences of the Group 1 (LB1-GR1), Group 2 (LB1-GR2), and Group 3 (LB1-GR3) LB1(f) peptides are indicated with arrows. The arrows encompass the LB1(f) peptides within the sequence of its natural context within the p5-like fimbrin protein. The LB1(f) polypeptide extends from Met1 to the C-terminal His residue before the stop codon.

20

**Figure 6:** Acrylamide gel stained with Coomassie showing the expression products of the following plasmids:

Lanes: 1. MW markers      2. pMGMCS      3. pRIT14588

4. LPD-LB1-A      5. LPD-LB1-II      6. LPD-LB1-III

25 7. LPD-LB1-III (LPD-LB1(f)<sub>2,1,3</sub> after purification process)    8. MW markers

**Figure 7:** Western Blot (using rabbit anti-LB1) of an acrylamide gel showing the expression products of the following plasmids:

Lanes: 1. MW markers      2. pMGMCS      3. pRIT14588

4. LPD-LB1-A      5. LPD-LB1-II      6. LPD-LB1-III  
 7. LPD-LB1-III (LPD-LB1(f)<sub>2,1,3</sub> after purification process)      8. MW markers

5      **Figure 8:** Western Blot (using a Mab anti-LPD) of an acrylamide gel showing the expression products of the following plasmids:

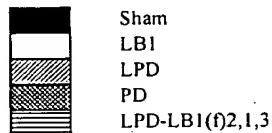
Lanes: 1. MW markers      2. pMGMCS      3. pRIT14588  
 4. LPD-LB1-A      5. LPD-LB1-II      6. LPD-LB1-III  
 7. LPD-LB1-III (LPD-LB1(f)<sub>2,1,3</sub> after purification process)      8. MW markers

10      **Figure 9:** Western Blot (using Ab against six-Histidine purification tag) of an acrylamide gel showing the expression products of the following plasmids:

Lanes: 1. MW markers      2. pMGMCS      3. pRIT14588  
 4. LPD-LB1-A      5. LPD-LB1-II      6. LPD-LB1-III  
 7. LPD-LB1-III (LPD-LB1(f)<sub>2,1,3</sub> after purification process)      8. MW markers

15      **Figure 10:** Passive transfer/challenge experiment. Mean tympanic membrane inflammation scores over the 35 day observation period for the 5 passively immunised chinchilla cohorts. The broken horizontal line at a mean tympanic membrane inflammation score of 1.5 indicates the level of inflammation attributable to adenovirus alone. Values above this line were considered to be an indication of ntHi-induced inflammation. ▼ - Sham; ○ - LB1; ■ - LPD; ◇ - PD; △ - LPD-LB1(f)<sub>2,1,3</sub>.

20      **Figure 11:** Bar graph showing the percentage of total middle ears known or suspected of containing an effusion based on otoscopy and tympanometry throughout the duration of the experiment. The time scale is measured with respect to the intranasal challenge of ntHi at day 0. Cohorts received antiserum directed against:



**DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT****Peptides of the Invention**

5 The peptides of the present invention relate to groups of newly identified LB1(f) peptides from equivalent regions of P5-like fimbrin proteins from various ntHi strains from Europe and the United States.

10 The DNA sequence of the P5-like fimbrin protein was ascertained from 83 strains of ntHi, and the peptide sequence of the LB1(f) peptide was noted. The peptide is a B-cell epitope which occurs in the same region (and within the same context) of each protein - approximately in the region that encompasses positions 110 and 140 of the amino acid sequence of the protein. In strain ntHi-10567RM, for example, the peptide exists between Arg117 to Gly135 (SEQ ID NO:1).

15 After alignment, the peptide sequences of both the American and European ntHi strains fell into the same three groups, with some variation within these groups. Group 1 peptides [or LB1(f)<sub>1</sub>] represented 71 % of the peptides, contained about 19 amino acids, and had not less than 75% identity with the peptide provided in SEQ ID NO:1. Group 2 peptides [or LB1(f)<sub>2</sub>] represented 19% of the peptides, contained 19-22 amino acids, and had not less than 75% identity with the peptide provided in SEQ ID NO:2. The group 20 could be additionally divided into 2 subgroups, group 2a [or LB1(f)<sub>2a</sub>] exemplified by SEQ ID NO:2, and group 2b [or LB1(f)<sub>2b</sub>] by SEQ ID NO:4. Group 3 peptides [or LB1(f)<sub>3</sub>] represented 10 % of the peptides, and contained 13 amino acids (provided in SEQ ID NO:3).

25 Of 83 ntHi strains analysed, the LB1(f) peptides from all 62 US strains and all 21 European strains fell into Groups 1-3. Table 1 shows all ntHi strains that were analysed and which Group their respective LB1(f) peptides belong to. Tables 2, 3, and 4 list the cumulated sequences of Group 1, 2, and 3 LB1(f) peptides respectively. Table 5 lists a representative example of a Group 1, 2a, 2b, and 3 LB1(f) peptide.

The previously known LB1(f) peptide (SEQ ID NO:5) falls into Group 1. Although it is known that this peptide is an effective immunogen, and confers protection against nTHi-caused *otitis media*, it has been unknown until now that this useful peptide exists in these three antigenically-distinct forms, which could be potentially combined to 5 provide protective immunogens against all *Haemophilus influenzae* strains that express the P5-like fimbrial protein.

The peptides of this invention relate to the representative peptides of Groups 1, 2a, 2b, and 3 (SEQ ID NO: 1, 2, 4, and 3 respectively), and to antigenically related variants of these peptides. Antigenically related variants can be either natural variants (as 10 exemplified by the peptides listed in tables 2, 3, and 4) or artificially modified variants that can be made by synthetic chemistry or recombinant DNA mutagenesis techniques that are well known to persons skilled in the art. The antigenically related variants of the peptides should have an amino acid sequence identity of at least 75 % to one of the peptides provided in SEQ ID NO:1-4 (and more preferably at least 85%, and most 15 preferably at least 95% identity), whilst still being capable of immunologically mimicking the corresponding antigenic determinant site of the P5-like fimbrial protein of non-typeable *Haemophilus influenzae*. For this invention "immunologically mimicking the corresponding antigenic determinant site of the P5-like fimbrial protein" is defined as the variant peptide being capable of inducing antibodies that recognise one of the wild- 20 type LB1(f) sequences (listed in tables 2, 3, and 4) in the context of the whole P5-like fimbrial protein AND/OR additionally defined as the variant peptide being capable of being recognised by the same antibody that recognises one of the wild-type LB1(f) sequences (listed in tables 2, 3, and 4) in the context of the whole P5-like fimbrial protein. In the first definition, the variant peptide should be capable of inducing such antibodies 25 either by itself, or in conjunction with a carrier molecule. In the second definition, the variant peptide should be capable of being recognised either by itself, or in conjunction with a carrier molecule. The antigenically related variant peptide does not include those peptides provided in SEQ ID NO: 5 (the previously determined LB1(f) peptide of P5-like

fimbrin protein from strain ntHi-1128) and SEQ ID NO:6 (the previously determined LB1(f)-like peptide of P5 protein from ntHi).

Antigenically related variants may have had amino acids added, inserted, substituted or deleted. Preferred variants are those that differ from the referents by 5 conservative amino acid substitutions.

The peptides of the invention also relates to combinations of LB1(f) peptides described above covalently linked, with optional spacer amino acids in between, to form a single peptide. For such peptides the peptides of SEQ ID NO: 5 & 6 can be used. The method to chemically synthesise or recombinantly express such peptides is well known 10 to a person skilled in the art. The optional spacer amino acids should preferably not be more than 18 amino acids either side of the peptide, and should preferably be composed of amino acids from the natural context of the LB1(f) peptide in the P5-like fimbrin protein (for example if two LB1(f) peptides were joined, the N-terminal LB1(f) peptide could have 9 amino acids of its natural C-terminal context linked to 9 amino acids of the 15 natural N-terminal context of the C-terminal LB1(f) peptide). One or more LB1(f) peptides may be linked in this way. Preferably 1-10 LB1(f) peptides are linked, more preferably 1-5, and still more preferably 1-3.

#### **Polypeptides of the Invention**

20 The polypeptides of the present invention relates to peptides described above being covalently linked to a carrier polypeptide that contains at least one T-cell epitope (for instance tetanus toxin, diphtheria toxin, OspA, P6 protein, OMP 26, protein D, or lipoprotein D) to form a chimeric LB1(f) polypeptide. This chimeric polypeptide comprises at least one of the peptides of the invention. Preferably the chimeric 25 polypeptide comprises 1-10 LB1(f) peptides, more preferably 1-5, and still more preferably 1-3. These peptides can be linked N-terminally, C-terminally, or both N- and C-terminally to the carrier polypeptide. Preferably, the carrier polypeptide is from *Haemophilus influenzae* so that it can act as a good immunogenic carrier, whilst having some protective efficacy in itself and/or whilst providing a source of homologous T-cell

epitopes derived from *H. influenzae*. Optionally, the chimeric polypeptide can also comprise a purification tag peptide sequence (such as a Histidine tag or a Glutathione-S-transferase tag) in order to aid subsequent purification of the polypeptide. Optional short peptide spacer sequences can be introduced between elements of the chimeric 5 polypeptide (as defined above in the Peptides of the Invention).

Preferably, the carrier polypeptide used is OMP26 of *H. influenzae* (WO 97/01638), or protein P6 of *H. influenzae* (Nelson, M. B. et al., (1988) Infection and Immunity 56, 128-134).

Most preferably, the carrier polypeptide used is protein D (PD) from non-typeable 10 *Haemophilus influenzae* or lipoprotein D (LPD - a lipidated form of PD). PD is a 42 kDa human IgD-binding outer surface protein that has been shown to be highly conserved among all strains of *Haemophilus influenzae* investigated so far (WO 91/18926). Both PD and LPD have been expressed in *E. coli*.

LPD was found to be a virulence factor in *H. influenzae*, and it elicits bactericidal 15 activity against ntHi in rat antisera. LPD from *H. influenzae* and the recombinantly-expressed equivalent of LPD can thus act as a good immunogenic carrier, whilst having some protective efficacy in itself. The non-lipidated form (PD) is more conveniently used for process reasons, and is also a potential carrier polypeptide of this invention. LPD is very immunogenic because of its built-in adjuvant properties; that is, its ability to induce 20 interleukins in macrophage and its ability to stimulate B cells to proliferate (WO 96/32963). PD does not have built-in adjuvant properties, and thus these conjugates require adjuvantation with, for example, aluminium hydroxide or aluminium phosphate. Antibody responses to LPD may protect against both typeable and nontypeable Hi strains. It thus represents an important carrier molecule for attaching other Hi antigens 25 (such as LB1(f) peptides) in order to obtain more effective vaccines against the organism. In addition to enhancing the immune response to the LB1(f) peptide antigen, LPD may serve as a protective antigen against both non-encapsulated and encapsulated strains of Hi.

Preferably three LB1(f) peptides are joined to the carrier polypeptide - one from each LB1(f) group. Preferably the LB1(f) peptides used are those provided in SEQ NO: 2, 3, and 5, and they are preferably linked C-terminally to the carrier polypeptide in the order SEQ ID NO: 2 (group 2 peptide), SEQ ID NO: 5 (group 1 peptide), SEQ ID NO: 3 (group 3 peptide). Such a polypeptide linked to LPD is known as LPD-LB1(f)<sub>2,1,3</sub>.

Although the chimeric polypeptide need not have a purification tag, when one is required a Histidine tag sequence is preferable, and it is preferably located at the C-terminus of the polypeptide.

The sequence of a preferred chimeric polypeptide is provided in Figure 5. 10 Residues 1-19 is the signal sequence of Protein D. This signal peptide may be removed in order to produce the PD version of the chimeric polypeptide.

Polypeptides of the present invention can be prepared in any suitable manner. Such polypeptides include recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for 15 preparing such polypeptides are well understood in the art, however examples of the method are presented in the Examples section.

### Polynucleotides of the Invention

The polynucleotides of the present invention relates to the wild-type 20 polynucleotide sequences of the LB1(f) peptides provided in Tables 6-8. They also relate to the wild-type DNA sequence of the polypeptides of the invention - that is to say constructing the chimeric polypeptide gene such that the wild-type gene sequence of the carrier polypeptide and wild-type polynucleotide sequences of LB1(f) peptides are used. Such a polynucleotide is provided in Figure 5. The DNA sequence of the optional spacer 25 amino acids is not essential for the invention, however where the spacer amino acids are from the natural context of the LB1(f) peptide, it is preferable (but not necessary) to use the natural DNA sequence of these spacers.

The polynucleotides of the invention also relates to DNA sequences that can be derived from the amino acid sequences of the peptides and polypeptides of the invention

bearing in mind the degeneracy of codons. This is well known in the art, as is knowledge of codon usage in different expression hosts which is helpful in optimising the recombinant expression of the peptides and polypeptides of the invention.

The invention also provides polynucleotides which are complementary to all the 5 above described polynucleotides.

When the polynucleotides of the invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself; or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a 10 leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions (for instance amino acid residues 1 to 19 of SEQ ID NO:7, the natural signal sequence of LPD). For example, a marker sequence which facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE 15 vector (Qiagen, Inc.) and described in Gentz *et al.*, *Proc Natl Acad Sci USA* (1989) 86:821-824, or is an HA tag, or is glutathione-s-transferase. Also preferred is LPD fused to its natural signal sequence (amino acid residues 1 to 19 of SEQ ID NO:7). The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that 20 stabilize mRNA.

#### **Vectors, Host Cells, Expression**

The present invention also relates to vectors which comprise a polynucleotide or polynucleotides of the present invention, and host cells which are genetically engineered 25 with vectors of the invention and to the production of peptides or polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate

expression systems or portions thereof for polynucleotides of the present invention. Introduction of polynucleotides into host cells can be effected by methods described in many standard laboratory manuals, such as Davis *et al.*, *BASIC METHODS IN MOLECULAR BIOLOGY* (1986) and Sambrook *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL*, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) such as calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, microinjection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as meningococci, streptococci, staphylococci, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used. Such systems include, among others, chromosomal, episomal and virus-derived systems, e.g., vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector suitable to maintain, propagate or express polynucleotides to produce a polypeptide in a host may be used. The appropriate nucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, *MOLECULAR CLONING, A LABORATORY MANUAL (supra)*.

For secretion of the translated protein into the lumen of the endoplasmic reticulum, into the periplasmic space or into the extracellular environment, appropriate secretion signals may be incorporated into the desired polypeptide. These signals may be

endogenous to the polypeptide (residues 1 to 19 of SEQ ID NO:7) or they may be heterologous signals.

#### **Purification of Recombinantly Expressed Peptides/Polypeptides**

5 Peptides and polypeptides of the invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulphate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most 10 preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during isolation and or purification.

15 Although the gene sequence of the chimeric LB1(f) polypeptide in the vector can be tagged with a Histidine-tag sequence which aids the purification of the polypeptide, it is not an essential element to the invention, as polypeptides without the Histidine-tag can still be purified by one of the techniques mentioned above.

20 Example 3 describes a purification method for purifying the LPD-LB1(f)(Group 2/Group 1/Group 3) (or LPD-LB1(f)<sub>2,1,3</sub>) chimeric polypeptide. A LPD-LB1(f) chimeric polypeptide with three or more LB1(f) peptides at the C-terminus of the polypeptide is easier to purify over one with only a single LB1(f) peptide at the C-terminus. This is due 25 to an observed partial degradation of the polypeptide from the C-terminus where it contains only one LB1(f) peptide that is not observed if there were three LB1(f) peptides at the C-terminus. Where some degradation has occurred, the full length polypeptide can be separated from the degraded form by incorporating a careful anion exchange step into the purification procedure.

#### **Antibodies**

The peptides and polypeptides of the invention, or cells expressing them can also be used as immunogens to produce antibodies immunospecific for the wild-type LB1(f)

peptides. The term "immunospecific" means that the antibodies have substantially greater affinity for the peptides or polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against the peptides or polypeptides can be obtained by 5 administering it to an animal, preferably a nonhuman, using routine protocols in the immunisation of an animal with an antigen, the collection of the blood, the isolation of the serum and the use of the antibodies that react with the peptide. The serum or IgG fraction containing the antibodies may be used in analysing the protein. For preparation of monoclonal antibodies, any technique which provides antibodies produced by 10 continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., *Nature* (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, *Immunology Today* (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, *MONOCLONAL ANTIBODIES AND CANCER THERAPY*, pp. 77-96, Alan R. Liss, Inc., 1985).

15 Techniques for the production of single chain antibodies (U.S. Patent No. 4,946,778) can also be adapted to produce single chain antibodies to peptides or polypeptides of this invention. Also, transgenic mice, or other organisms including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones 20 expressing the peptide or to purify the peptides or polypeptides of the invention by affinity chromatography.

The peptides and polypeptides of the present invention also are useful to produce polyclonal antibodies for use in passive immunotherapy against *H. influenzae*. Human immunoglobulin is preferred because heterologous immunoglobulin may provoke a 25 deleterious immune response to its foreign immunogenic components. Polyclonal antisera is obtained from individuals immunized with the peptides or polypeptides in any of the forms described. The immunoglobulin fraction is then enriched. For example, immunoglobulins specific for epitopes of the protein are enriched by immunoaffinity techniques employing the peptides or polypeptides of this invention. The antibody is

specifically absorbed from antisera onto an immunoadsorbent containing epitopes of the polypeptide and then eluted from the immunoadsorbent as an enriched fraction of immunoglobulin.

## 5   Vaccines

The earlier work on the LB1(f) peptide from strain ntHi-1128 indicated that this peptide could be used as an immunogen for the development of a subunit vaccine against *Haemophilus influenzae* disease, particularly to prevent or reduce susceptibility to acute *otitis media* and other diseases caused by nontypeable strains. This invention extends this work by identifying three main Groups of LB1(f) peptides. The differences between the three groups were such that it is unlikely that a cross protection could be achieved between strains belonging to different groups. The use of examples from each of these peptide groups in combination is therefore able to provide a more efficient and complete vaccine against strains of *Haemophilus influenzae* (preferably ntHi) that express the P5-like fimbrin protein.

Another aspect of the invention therefore relates to a method for inducing an immunological response in a mammal which comprises inoculating the mammal with a peptide or polypeptide of the invention adequate to produce antibody and/or T cell immune response to protect said animal from *H. influenzae* disease, among others. Yet another aspect of the invention relates to a method of inducing immunological response in a mammal which comprises, delivering a peptide or polypeptide of the invention via a vector directing expression of a polynucleotide of the invention *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases.

25       A further aspect of the invention relates to an immunological/vaccine formulation (composition) which, when introduced into a mammalian host, induces an immunological response in that mammal to a LB1(f) peptide or polypeptide wherein the composition comprises a LB1(f) peptide or polypeptide gene, or LB1(f) peptide or polypeptide itself. The vaccine formulation may further comprise a suitable carrier. The

LB1(f) vaccine composition is preferably administered orally, intranasally or parenterally (including subcutaneous, intramuscular, intravenous, intradermal, transdermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use.

10 The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

15 Yet another aspect relates to an immunological/vaccine formulation which comprises the polynucleotide of the invention. Such techniques are known in the art, see for example Wolff *et al.*, *Science*, (1990) 247: 1465-8.

20 The peptides or polypeptides of this invention can be administered as multivalent subunit vaccines in combination with antigens from other proteins of *H. influenzae* to achieve an enhanced bactericidal activity. For combined administration with epitopes of other proteins, the LB1(f) peptide or polypeptide is either administered separately, as a mixture or as a conjugate or genetic fusion polypeptide. The conjugate is formed by standard techniques for coupling proteinaceous materials. The peptides or polypeptides of the invention can be used in conjunction with antigens of other organisms (e.g. encapsulated or nonencapsulated, bacteria, viruses, fungi and parasites). For example, the 25 peptides or polypeptides of the invention are useful in conjunction with antigens of other microorganisms implicated in *otitis media* or other diseases. These include *Streptococcus pneumoniae*, *Streptococcus pyogenes*, group A, *Staphylococcus aureus*, respiratory syncytial virus and *Branhemella catarrhalis*.

The evaluation of the peptides or polypeptides of the invention as potential vaccines against ntHi-caused *otitis media* is made in a chinchilla animal model developed by Dr. L. Bakaletz of Ohio State University. This model mimics the development of *otitis media* in children and is based on the successive intranasal 5 administrations of adenovirus and ntHi a week apart. In these conditions, the bacteria is able, after the colonisation of the nasopharynx, to invade the middle ear via the Eustachian tube. Once there, ntHi will proliferate and induce an inflammatory process similar to what is observed in children.

For vaccine evaluation, by the time the chinchilla has been actively immunised 10 they are too old at the time of challenge to be inoculated by the intranasal route with ntHi: even with a preinfection with adenovirus, almost none of them will develop *otitis media*. As an alternative route of challenge, a direct inoculation of the bacteria into the middle ear (bullae) through the skull is used. Passive transfer/challenge protocols can also be used to avoid needing trans-bullar challenge.

15 With all these types of challenge, the severity of the disease can be scored by otoscopic observation (through the external ear) or tympanometry, which evaluate the level of inflammation in the middle ear or changes in middle ear pressure and presence of fluid in the middle ear, respectively. The efficacy of a vaccine is determined by the reduction of the severity and/or the duration of the inflammation and the reduction of the 20 colonisation in the ear and the nasopharynx.

In previous experiments, the protective efficacy of both LB1 from strain ntHi- 25 1128 and LPD was evaluated after active immunisation, and intrabullar challenge. Repeatedly, immunisation with LB1 protected chinchilla against *otitis media* as indicated by a reduced length of *otitis*, reduced severity, and reduced colonisation in both the ears and the nasopharynx. The immunisation with LPD protected chinchillas against *otitis media* but not as well as LB1, and not reproducibly.

The vaccines of the invention can be further evaluated by examining whether the peptides or polypeptides of the invention inhibit adherence of ntHi to chinchilla epithelial throat cells, and whether they can prevent nasopharyngeal colonisation by ntHi *in vivo*.

The LB1 peptide from ntHi-1128 has a dose-dependent effect on the inhibition of the adherence of ntHi to chinchilla epithelial throat cells (probably as it acts as a direct steric inhibitor of ntHi binding), and lowers the ntHi in nasopharyngeal lavage fluid. Nasopharyngeal colonisation is an initial step required for the development of *otitis media*, 5 therefore this inhibition of colonisation will also help to inhibit the development of *otitis media*.

### Diagnostic Assays/Kits

This invention also relates to the use of the peptides or polypeptides of the 10 invention, and antibodies against these peptides or polypeptides as diagnostic reagents. Detection of a LB1(f) peptide will provide a diagnostic tool that can add to or define a diagnosis of *Haemophilus influenzae* disease, among others.

Biological samples for diagnosis may be obtained from a subject's cells, such as from serum, blood, urine, saliva, tissue biopsy, sputum, lavage fluids.

15 Polynucleotides of the invention, which are identical or sufficiently identical to one of the nucleotide sequences contained in Tables 6-8, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification (PCR) reaction, to isolate full-length cDNAs and genomic clones encoding P5-like fimbrial protein. Such hybridization techniques are known to those of skill in the art. Typically these nucleotide 20 sequences are 80% identical, preferably 90% identical, more preferably 95% identical to that of the referent. The probes generally will comprise at least 15 nucleotides. Preferably, such probes will have at least 30 nucleotides and may have at least 50 nucleotides. Particularly preferred probes will range between 30 and 50 nucleotides. In this way *Haemophilus influenzae* can be detected in a biological sample, and under particularly 25 stringent hybridisation conditions, the specific strain or strains of *Haemophilus influenzae* present in a sample could be ascertained using the wild-type polynucleotide sequences provided in Tables 6-8.

Thus in another aspect, the present invention relates to a diagnostic kit for a disease, particularly *Haemophilus influenzae* disease, which comprises:

(a) a polynucleotide of the invention, preferably a nucleotide sequence provided in Tables 6-8;

(b) a nucleotide sequence complementary to that of (a);

(c) a LB1(f) peptide of the invention, preferably the peptides of SEQ ID NO: 1-4; or

5 (d) an antibody to a LB1(f) peptide of the invention, preferably to the peptides of SEQ ID NO: 1-4.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

10 The invention is further illustrated by the following examples.

15 **Examples**

The examples below are carried out using standard techniques, which are well known and routine to those skilled in the art, except where otherwise described in detail.

The examples illustrate, but do not limit the invention.

20 Example 1: The determination of the amino acid sequence variability of the LB1(f) peptide in various ntHi strains.

1a) Culture of ntHi isolates - the preparation of samples for PCR analysis

53 ntHi isolates were obtained from Dr. L. Bakaletz of Ohio State University, and 30 ntHi isolates were obtained from Dr. A. Forsgren of Malmö, Sweden.

25 0.1 mL of a liquid culture of each ntHi isolate was spread on Gelose Chocolate Agar (GCA). The purity of the samples was controlled on solidified media (TSA - Tryptose Soy Agar in Petri dishes). The dishes were incubated at 35°C for 24 hours. Colonies from dishes were resuspended in 5 mL of filtered TSB (Tryptose Soy Broth + 3 µg/µl NAD; + 3 µg/µl Hémine, + 1% horse serum). 50 mL of TSB liquid media was

inoculated with 2.5 mL of the culture, and were incubated at 35°C. When the concentration of the culture grew to 10<sup>8</sup> cells/mL, 10 mL of culture were centrifuged at 10,000 rpm, 4°C for 15 minutes. The supernatant was removed and the cells were washed in physiological buffer. The cells were centrifuged at 10,000 rpm for 15 minutes, 5 4°C. The cells were resuspended at a final concentration of 10<sup>9</sup> cells /mL. The cells were boiled at 95 -100°C for 10 - 15 minutes, and then placed directly on ice. Samples were frozen at -70°C. The samples were then ready for DNA amplification by PCR.

1b) Amplification of P5-like fimbrin gene DNA fragment by PCR

10 PCR amplification of fragment of the fimbrin gene were performed on the ntHi preparations from example 1a). 200 µL of an ntHi preparation were centrifuged 14,200 rpm for 3 minutes at room temperature. All the supernatant was removed. The cells were resuspended in 25 µL of ADI, were boiled at 95°C for 10 minutes, and were centrifuge for 3 minutes at 14,200 rpm. 5 µL of supernatant were used for a PCR reaction.

15 Amplification of DNA was performed with specific primers:

NTHi-01: - 5' - ACT-GCA-ATC-GCA-TTA-GTA-GTT-GC - 3'

NTHi-02: - 5' - CCA-AAT-GCG-AAA-GTT-ACA-TCA-G - 3'

20 The PCR reaction mixture was composed of the following: cell extract supernatant, 5.0 µL; Primer NTHi-01 (1/10), 1.0 µL; Primer NTHi-02 (1/10), 1.0 µL; DMSO, 2.0 µL; dNTP mix, 4.0 µL; Buffer 10x, 5.0 µL; ADI, 31.5 µL; *Taq* polymerase, 0.5 µL.

The PCR cycle conditions were as follows: (94°C for 1 min; 50°C for 1 min; 72°C for 3 min) for 25 cycles, and finishing with 72°C for 10 min. The reaction was monitored by electrophoresis in a 3% agarose gel in TBE buffer.

25 The primers used for the identification of which group a particular ntHi P5-like fimbrin LB1(f) peptide belonged to are as follows (they are used in a similar way to the reaction above):

Group 1:

NTHi-01 : 5'-ACT-GCA-ATC-GCA-TTA-GTA-GTT-GC-3'

NTHi-GR1 : 5'-GTG-GTC-ACG-AGT-ACC-G-3'  
Group 2:  
NTHi-01 : 5'-ACT-GCA-ATC-GCA-TTA-GTA-GTT-GC-3'  
NTHi-GR2bis : 5'-TCT-GTG-ATG-TTC-GCC-TAG-3'  
5 Group 3:  
NTHi-01 : 5'-ACT-GCA-ATC-GCA-TTA-GTA-GTT-GC-3'  
NTHi-GR3 : 5'-CTA-TCG-ATG-CGT-TTA-TTA-TC-3'

1c) DNA purification

10 The PCR Clean Up Kit for purification of DNA fragments from PCR reactions was used (Boehringer Mannheim). At the end of the procedure, the purified PCR product was eluted twice in 25  $\mu$ L volumes of redistilled water from the silica resin.

The purified products were analyzed by electrophoresis in a 3% agarose gel stained with ethidium bromide. The DNA was then ready for sequencing .

15

1d) DNA sequencing

This was done using an ABI Automatic Sequence, the ABI-PRISM - DNA sequencing Kit (using Terminator PCR Cycle Sequencing), and AmpliTaq DNA Polymerase FS (from Perkin Elmer).

20

The PCR reaction mixture used was as follows: Mix (from the kit), 8.0  $\mu$ L; DNA (approx. 1  $\mu$ g), 3.0  $\mu$ L; Primer (see below) 1/5 or 1/10, 1.0  $\mu$ L; ADI, 8.0  $\mu$ L

The sequencing primers used were as follows:

NTHi -03: 5'-AGG-TTA-CGA-CGA-TTT-CGG-3' or

NTHi -04: 5'- CGC-GAG-TTA-GCC-ATT-GG-3' or

25

NTHi -05: 5'- AAA-GCA-GGT-GCT-TTA-G-3' or

NTHi -06: 5'- TAC-TGC-GTA-TTC-TGC-ACC-3'

OR

NTHi-03: 5'-AGG-TTA-CGA-CGA-TTT-CGG- 3'

NTHi-04: 5'-CGC-GAG-TTA-GCC-ATT-GG-3'

NTHi-05: 5'-AAA-GCA-GGT-GTT-GCT-TTA-G-3'  
NTHi-06: 5'-TAC-TGC-GTA-TTC-TTA-TGC-ACC-3'  
NTHi-14: 5'-GGT-GTA-TTT-GGT-GGT-TAC-C-3'  
NTHi-15: 5'-GTT-ACG-ACG-ATT-ACG-GTC-G-3'

5

The PCR cycle sequencing conditions were as follows: (96°C for 30 seconds; 50°C for 15 seconds; 60°C for 4 min) for 25 cycles, and finishing with 72°C for 10 min.

The PCR product was prepared and analysed by: adding 80 µL ADI to the PCR sequence reaction to obtain a final volume of 100 µL; adding an equal volume of phenol/chloroform to the DNA solution. The sample was then centrifuged at 14,500 rpm at 4°C for 3 min and the top aqueous layer was removed. The phenol/chloroform step and the centrifugation step were repeated once more. 10 µL 3M NaAc pH 4.8 and 220 µL 100% ethanol (at room temperature) were then added and mixed. The sample was placed at -20°C for 5 min, and then centrifuged at 14,000 rpm 20 min at 4°C. The ethanol supernatant was removed and the pellet was rinsed with 1 mL of 70% ethanol (at room temperature). This was centrifuged at 14,000 rpm 10 min at 4°C, and the supernatant was removed as before. The pellet was air dried, and frozen overnight. The pellet was dissolved in the following solution: formamide 100% deionised water, 5 volumes; 0.5M EDTA pH 8.00, 1 volume. This was vortexed a few seconds and loaded on a sequencing 10 15 20 gel.

#### 1e) Cumulated results and conclusions

A list of the various ntHi isolates that were analysed in terms of the sequence of their LB1(f) peptides from P5-like fimbrial protein is shown in Table 1. The group classification was determined by aligning the LB1(f) peptide against SEQ ID NO: 5, 2, or 3 (being the representative Group 1, 2 or 3 LB1(f) peptides respectively). LB1(f) peptides had to have at least 75% identity with the representative peptide of a group in order for the classification of the group to be assigned to the test peptide. Tables 2, 3, and 4 show the aligned sequences of the Group 1, 2, and 3 LB1(f) peptide sequences 25

respectively. Table 5 shows the representative LB1(f) peptides of Group 1, 2a, 2b, and 3 aligned with respect to each other.

Tables 6-9 show the DNA sequences of the LB1(f) peptides of Tables 2-5, respectively.

Table: 1

	Serotype	n°order	Strains	Group
1	NTHi	1848L	H. influenzae	1
2	NTHi	1848NP	H. influenzae	1
3	NTHi	1885R	H. influenzae	1
4	NTHi	1885MEE	H. influenzae	2
5	NTHi	10547RMEE	H. influenzae	3
6	NTHi	10548LMEE	H. influenzae	3
7	NTHi	10567RMEE	H. influenzae	1
8	NTHi	10568LMEE	H. influenzae	1
9	NTHi	10567&8NP	H. influenzae	3
10	NTHi	1371MEE	H. influenzae	1
11	NTHi	214NP	H. influenzae	1
12	NTHi	1370MEE	H. influenzae	1
13	NTHi	1380MEE	H. influenzae	1
14	NTHi	217NP	H. influenzae	1
15	NTHi	266NP	H. influenzae	2
16	NTHi	167NP	H. influenzae	1
17	NTHi	1657MEE	H. influenzae	1
18	NTHi	284NP	H. influenzae	1
19	NTHi	1666MEE	H. influenzae	1
20	NTHi	287NP	H. influenzae	1
21	NTHi	1236MEE	H. influenzae	2
22	NTHi	183NP	H. influenzae	2
23	NTHi	165NP	H. influenzae	2
24	NTHi	1182MEE	H. influenzae	1
25	NTHi	166NP	H. influenzae	1
26	NTHi	1199MEE	H. influenzae	1
27	NTHi	172NP	H. influenzae	1
28	NTHi	1230MEE	H. influenzae	1
29	NTHi	180NP	H. influenzae	1
30	NTHi	1234MEE	H. influenzae	1
31	NTHi	182NP	H. influenzae	1
32	NTHi	152NP	H. influenzae	1
33	NTHi	226NP	H. influenzae	1
34	NTHi	1714MEE	H. influenzae	2
35	NTHi	297NP	H. influenzae	2
36	NTHi	1715MEE	H. influenzae	2

37	NTHi	1729MEE	H. influenzae	3
38	NTHi	1728MEE	H. influenzae	3
39	NTHi	250NP	H. influenzae	1
40	NTHi	1563MEE	H. influenzae	1
41	NTHi	1562MEE	H. influenzae	1
42	NTHi	10559RMEE	H. influenzae	1
43	NTHi	1712MEE	H. influenzae	1
44	NTHi	1521	H. influenzae	1
45	NTHi	1060RMEE	H. influenzae	1
46	NTHi	86-027MEE	H. influenzae	2
47	NTHi	86-027NP	H. influenzae	1
48	NTHi	86-028NP	H. influenzae	1
49	NTHi	86-028LMEE	H. influenzae	1
50	NTHi	90-100	H. influenzae	1
51	NTHi	90-121RMEE	H. influenzae	1
52	NTHi	1128	H. influenzae	1
53	NTHi	90-100RMEE	H. influenzae	1
54	NTHi*	476	H. influenzae	1
55	NTHi*	480	H. influenzae	1
56	NTHi*	481	H. influenzae	1
57	NTHi*	482	H. influenzae	1
58	NTHi*	484	H. influenzae	1
59	NTHi*	486	H. influenzae	1
60	NTHi*	490	H. influenzae	1
61	NTHi*	492	H. influenzae	2
62	NTHi*	494	H. influenzae	1
63	NTHi*	495	H. influenzae	2
64	NTHi*	498	H. influenzae	1
65	NTHi*	499	H. influenzae	1
66	NTHi*	500	H. influenzae	2
67	NTHi*	501	H. influenzae	1
68	NTHi*	502	H. influenzae	2
69	NTHi*	503	H. influenzae	1
70	NTHi*	504	H. influenzae	3
71	NTHi*	506	H. influenzae	2
72	NTHi*	507	H. influenzae	1
73	NTHi*	546	H. influenzae	2
74	NTHi*	567	H. influenzae	1
75	NTHi	544	H. influenzae	3

76	NTHi	565	H. influenzae	1	
77	NTHi	600	H. influenzae		3
78	NTHi	601	H. influenzae	1	
79	NTHi	603	H. influenzae	1	
80	NTHi	604	H. influenzae		2
81	NTHi	605	H. influenzae	1	
82	NTHi	606	H. influenzae	1	
83	NTHi	608	H. influenzae	1	

Cumulated list of ntHi strains investigated and the classification of the sequence of their respective LB1(f) peptides from P5-like fimbrin protein (strains 1-53 from L. Bakalcz, strains 54-83 from A. Forsgren). \* denotes a European strain of ntHi, all others were isolated from the United States. Strains 1885 and 1128 are available from the American

5 Type Culture Collection (ATCC # 55431 and 55430 respectively).

Table: 2 - Cumulated Group 1 Peptide Sequences

N1128	RSDYKFYEDANGTRDHKKG
N1380MEE	RSDYKFYEDANGTRDHKKG
N1885R	RSDYKFYEDANGTRDHKKG
N1562MEE	RSDYKFYEDANGTRDHKKG
N1563MEE	RSDYKFYEDANGTRDHKKG
N180NP	RSDYKFYEDANGTRDHKKG
N217NP	RSDYKFYEDANGTRDHKKG
N284NP	RSDYKFYEDANGTRDHKKG
N1666MEE	RSDYKFYEDANGTRDHKKG
N1230MEE	RSDYKFYEDANGTRDHKKG
NTHI-501	RSDYKFYEDANGTRDHKKG
NTHI-507	RSDYKFYEDANGTRDHKKG
NTHI-565	RSDYKFYEDANGTRDHKKG
NTHI-603	RSDYKFYEDANGTRDHKKG
NTHI-608	RSDYKFYEDANGTRDHKKG
N287NP	RSDYKFYEDANGTRDHKKG
N86028LM	RSDYKFYEDANGTRDHKKG
N86028NP	RSDYKFYEDANGTRDHKKG
N152NP	RSDYKFYEDADGTRDHKKG
N1234MEE	RSDYKFYDDANGTRDHKKG
N182NP	RSDYKFYDDANGTRDHKKG
N90100RM	RSDYKFYEDENGTRDHKKG
N90100	RSDYKFYEDENGTRDHKKG
N10567RM	RSDYKFYEAAANGTRDHKKG
N1060MEE	RSDYKFYEAAANGTRDHKKG
N172NP	RSDYKFYEAAANGTRDHKKG
N1199MEE	RSDYKFYEAAANGTRDHKKG
N10568LM	RSDYKFYEAAANGTRDHKKG
N90121RM	RSDYKFYEAAANGTRDHKKG
N86027NP	RSDYKFYEVANGTRDHKKG
NTHI-486	RSDYKFYEVANGTRDHKKG
N1712MEE	RSDYKFYEVANGTRDHKKG
NTHI-503	RSDYKFYEAAANGTRDHKKG
NTHI-476	RSDYKFYEEAANGTRDHKKG
N166NP	RSDYKFYNDANGTRDHKKKS
N1182MEE	RSDYKFYNDANGTRDHKKKS
N1848NP	RSDYKFYEVANGTRDHKKKS
N1371MEE	RSDYKFYEVANGTRDHKKKS
NTHI-498	RSDYKFYEVANGTRDHKKKS
NTHI-606	RSDYKFYEVANGTRDHKKKS
N1848L	RSDYKFYEVANGTRDHKKKS
NTHI-567	RSDYKFYEDANGTRDRKTG
NTHI-484	RSDYKFYEDANGTRKHKEG
N10559RM	RSDYKLYEVANGTRDHKKKS
NTHI-601	RSDYKFYEVANGTRDHKQS
NTHI-481	RSDYKFYEVANGTRDHKQS
NTHI-482	RSDYKFYEVANGTRDHKQS
N1370MEE	RSDYKFYEVANGTRDHKQS
N226NP	RSDYKFYEEAANGTRDHKRS
NTHI-480	RSDYKFYEDANGTRERKRG

N1657MEE	RSDYKFYEVANGTRERKKG
N267NP	RSDYKFYEVANGTRERKKG
NTHI-490	RSDYKFYEVANGTRERKKG
NTHI-494	RSDYKFYEVANGTRERKKG
N214NP	RSDYKFYEVNGTRDHKQS
N250NP	RSDYKRYEEANGTRNHDKG
N1521	RSDYKRYEEANGTRNHDKG
NTHI-605	RSDYKRYEEANGTRNHDKG
NTHI-499	RSDYEFYEAPNSTRDHKKKG

**Table: 3 - Cumulated Group 2 Peptide Sequences**

N1715MEE	RSDYKLYNKNSSSNSTLKNLGE
N1714MEE	RSDYKLYNKNSSSNSTLKNLGE
N86027RM	RSDYKLYNKNSSSNSTLKNLGE
N297NP	RSDYKLYNKNSSSNSTLKNLGE
N266NP	RSDYKLYNKNSSSNSTLKNLGE
N1885MEE	RSDYKLYNKNSSSNSTLKNLGE
NTHI-546	RSDYKLYNKNSSSNSTLKNLGE
NTHI-604	RSDYKLYNKNSSSNSTLKNLGE
NTHI-492	RSDYKLYNKNS -NSTLKNLGE
NTHI-502	RSDYKLYDKNSSS -TLKKLGE
NTHI-506	RSDYKLYNKNS -NSTLKNLGE
N1236MEE	RSDYKLYNKNS --- TLKDLGE
NTHI-500	RSDYKLYNKNS --- TLKDLGE
NTHI-183	RSDYKLYNKNS --- TLKDLGE
N165NP	RSDYKLYNKNSN -- TLKDLGE
NTHI-495	RSDYKLYNKNSSD -- ALKKLGE

5

**Table: 4- Cumulated Group 3 Peptide Sequences**

N1729MEE	RSDYKFYDNKRID
NTHI-504	RSDYKFYDNKRID
NTHI-544	RSDYKFYDNKRID
NTHI-600	RSDYKFYDNKRID
N1728MEE	RSDYKFYDNKRID
N10548LM	RSDYKFYDNKRID
N10547RM	RSDYKFYDNKRID
N105678R	RSDYKFYDNKRID

10

**Table: 5 - Cumulated Group 1, 2a, 2b, and 3 Peptide Sequences**

N1128	RSDYKFYEDANGTRDHKKG---
N1715MEE	RSDYKLYNKNSSSNSTLKNLGE
NTHI-183	RSDYKLYNKNS --- TLKDLGE
N1729MEE	RSDYKFYDN-----KRID---

**Table: 6 - Cumulated Group 1 Gene Sequences**

N1128	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1380MEE	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1885R	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1562MEE	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1563MEE	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N180NP	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N217NP	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N284NP	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1666MEE	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N1230MEE	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-501	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-507	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-565	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-603	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-608	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N287NP	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N86028LM	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N86028NP	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGAAAGGT
N152NP	CGTTCTGATTATAAATTTATGAAGATGCAGACGGTACTCGTGACCACAAGAAAGGT
N1234MEE	CGTTCTGATTATAAATTTATGATGATGCAACGGTACTCGTGACCACAAGAAAGGT
182NP	CGTTCTGATTATAAATTTATGATGATGCAACGGTACTCGTGACCACAAGAAAGGT
N90100RM	CGTTCTGATTATAAATTTATGAAGATGAAAACGGTACTCGTGACCACAAGAAAGGT
N90100	CGTTCTGATTATAAATTTATGAAGATGAAAACGGTACTCGTGACCACAAGAAAGGT
N10567RM	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N1060MEE	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N172NP	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N1199MEE	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N10568LM	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N90121RM	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
N86027NP	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-486	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAGGT
N1712MEE	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-503	CGTTCTGATTATAAATTTATGAAGCTGCAAACGGTACTCGTGACCACAAGAAAGGT
NTHI-476	CGTTCTGATTATAAATTTATGAAGAGGCAAACGGTACTCGTGACCACAAGAAAGGT
N166NP	CGTTCTGATTATAAATTTATAATGATGCAACGGTACTCGTGACCACAAGAAAAGT
N1182MEE	CGTTCTGATTATAAATTTATAATGATGCAACGGTACTCGTGACCACAAGAAAAGT
N1848NP	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
N1371MEE	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
NTHI-498	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
NTHI-606	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
N1848L	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
NTHI-567	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGACCACAAGACAGGT
NTHI-484	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTAAGCACAAGGAAGGT
N10559RM	CGTTCTGATTATAAACTTATGAAGTTGCAAACGGTACTCGTGACCACAAGAAAAGT
NTHI-601	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGCAAAGT
NTHI-481	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGCAAAGT
NTHI-482	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGCAAAGT
N1370MEE	CGTTCTGATTATAAATTTATGAAGTTGCAAACGGTACTCGTGACCACAAGCAAAGT
N226NP	CGTTCTGATTATAAATTTATGAAGAGGCAAACGGTACTCGTGACCACAAGAGAGGT
NTHI-480	CGTTCTGATTATAAATTTATGAAGATGCAAACGGTACTCGTGAGCGCAAGAGAGGT

N1657MEE	CGTTCTGATTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N267NP	CGTTCTGATTATAAATTGTACAATAAAACGGTACTCGTGAGCGCAAGAAAGGT
NTHI-490	CGTTCTGATTATAAATTGTACAATAAAACGGTACTCGTGAGCGCAAGAAAGGT
NTHI-494	CGTTCTGATTATAAATTGTACAATAAAACGGTACTCGTGAGCGCAAGAAAGGT
N214NP	CGTTCTGATTATAAATTGTACAATAACGGTACTCGTGACCAAGCAAAGT
N250NP	CGTTCTGATTATAAACGTTATGAAGAAGCAAACGGTACTCGTAACCACGACAAAGGT
N1521	CGTTCTGATTATAAACGTTATGAAGAAGCAAACGGTACTCGTAACCACGACAAAGGT
NTHI-605	CGTTCTGATTATAAACGTTATGAAGAAGCAAACGGTACTCGTAACCACGACAAAGGT
NTHI-499	CGTTCTGATTATAAACGTTATGAAGCTCCAAACAGTACTCGTGACCAAGCAAAGGT

**Table: 7 - Cumulated Group 2 Gene Sequences**

N1715MEE	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N1714MEE	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N86027RM	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N297NP	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N266NP	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N1885MEE	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
NTHI-546	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
NTHI-604	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
NTHI-492	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---AATAGTACTCTTAAAAACCTAGGC
NTHI-502	CGTTCTGACTATAAATTGTACGATAAAATAGTAGTAGTAAT---ACTCTTAAAAACCTAGGC
NTHI-506	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---AATAGTACTCTTAAAAACCTAGGC
N1236MEE	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---ACTCTTAAAGACCTAGGC
NTHI-500	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---ACTCTTAAAGACCTAGGC
NTHI-183	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---ACTCTTAAAGACCTAGGC
N165NP	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAAT---ACTCTTAAAGACCTAGGC
NTHI-495	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTGAT---GCTCTTAAAAACCTAGGC

5

**Table: 8 - Cumulated Group 3 Gene Sequences**

N1729MEE	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
NTHI-504	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
NTHI-544	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
NTHI-600	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
N1728MEE	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
N10548LM	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
N10547RM	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT
N105678R	CGTTCTGACTATAAATTCTACGATAATAAACGCATCGAT

10

**Table: 9 - Cumulated Group 1, 2a, 2b, and 3 Gene Sequences**

N1128	CGTTCTGATTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
N1715MEE	CGTTCTGACTATAAATTGTACAATAAAATAGTAGTAGTAATAGTACTCTTAAAAACCTAGGC
NTHI-183	CGTTCTGACTATAAATTGTACAATAAAATAGTAGT---ACTCTTAAAGACCTAGGC
N1729MEE	CGTTCTGACTATAAATTCTACGATAAT---AACGCATCGAT

The study shows that the LB1(f) peptides of the P5-like fimbrin protein from all 83 ntHi isolates tested can be classified in three groups, and that both United States and European ntHi isolates fall into this classification.

5

Example 2: The Expression of LPD-LB1(f) peptide fusion polypeptides in *E. coli*

#### Source Material

##### 1) The expression vector pMG1

The expression vector pMG1 is a derivative of pBR322 in which bacteriophage  $\lambda$  derived control elements for transcription and translation of foreign inserted genes were introduced (Young et al. (1983) PNAS USA 80, 6105-6109). In addition, the Ampicillin resistance gene was exchanged with the Kanamycin resistance gene.

The vector contains the  $\lambda$  promoter  $P_L$ , operator  $O_L$  and two utilization sites ( $Nu_t_L$  and  $Nu_t_R$ ) to relieve transcriptional polarity effects. Vectors containing the  $P_L$  promoter, 15 are introduced into an *E. coli* lysogenic host to stabilize the plasmid DNA. Lysogenic host strains contain replication-defective  $\lambda$  phage DNA integrated into the genome. The chromosomal  $\lambda$  phage DNA directs the synthesis of the  $cI$  repressor protein which binds to the  $O_L$  repressor of the vector and prevents binding of RNA polymerase to the  $P_L$  promoter and thereby transcription of the inserted gene. The  $cI$  gene of the expression 20 strain AR58 contains a temperature sensitive mutant so that  $P_L$  directed transcription can be regulated by temperature shift, i.e. an increase in culture temperature inactivates the repressor and synthesis of the foreign protein is initiated. This expression system allows controlled synthesis of foreign proteins especially of those that may be toxic to the cell.

25 2) The expression vector pMGMCS

The nucleotide sequence between the BamHI and the XbaI restriction sites in pMG1 was replaced by a multiple cloning site DNA fragment (MCS) to generate the pMGMCS expression vector (fig. 1).

A poly-His sequence has been added at the 3' end of the MCS sequence to allow the expression of a protein product fused to a 6-Histidine tail.

The sequence coding for the first 3 amino acids of NS1 (Met-Asp-Pro) is present on the vector, before the BamHI restriction site.

5

### 3) Construction of vector pRIT14588

The cloning strategy for the generation of the pRIT14588 expression vector from the pMGMCS vector is outlined in Fig. 2. The lipoprotein D gene was amplified by PCR from the pHIC348 vector (Janson et al. (1991) Infect. Immun. 59, 119-125) with PCR primers containing BamHI and NcoI restriction sites at the 5' and 3' ends, respectively. The BamHI/NcoI fragment was then introduced into pMGMCS between BamHI and NcoI.

The lipoprotein D gene product contains its native signal sequence except for the first three amino acids which have been replaced by Met-Asp-Pro from NS1.

15 pRIT14588 was used to introduce LB1(f) peptides to the 3' end of the Lipoprotein D gene. The LB1(f) peptides used were the following: group 1, ntHi-1128 (SEQ ID NO:5); group 2, ntHi-1715 MEE (SEQ ID NO: 2); group 3, ntHi-1729 MEE (SEQ ID NO: 3).

20

### 4) The *E. coli* strain AR58

The AR58 lysogenic *E. coli* strain used for the production of the protein D carrier protein is a derivative of the standard NIH *E. coli* K12 strain N99 ( $F^+$   $su^+$   $galK2^+$ ,  $lacZ^+$   $thr^+$ ). It contains a defective lysogenic  $\lambda$  phage ( $galE::TN10$ ,  $\lambda$   $Kil^+$   $cl857^+$   $DH1$ ). The  $Kil^+$  phenotype prevents the shut down of host macromolecular synthesis. The  $cl857^+$  mutation confers a temperature sensitive lesion to the  $cl$  repressor. The  $DH1$  deletion removes the  $\lambda$  phage right operon and the hosts *bio*, *uvr3*, and *chlA* loci. The AR58 strain (Mott et al. (1985) PNAS USA. 82, 88-92) was generated by transduction of N99 with a P1 phage stock previously grown on an SA500 derivative ( $galE::TN10$ ,  $\lambda$ ,  $Kil^+$   $cl857^+$   $DH1$ ). The introduction of the defective lysogen into N99 was selected with

tetracycline (a TN10 transposon coding for tetracyclin resistance is present in the adjacent galE gene).

Example 2a) Producing a Lipoprotein D - LB1(f) Group 1 fusion

5 The aim of this construct was to clone the 19 residue LB1(f) peptide 3' to the *NcoI* site of the multiple cloning site of pRIT14588. Immediately 3' to the *NcoI* site, two Glycine residues were introduced to place the LB1(f) peptide gene in frame with the LPD gene. After the two Gly residues, the DNA coding for 8 natural residues N-terminal to the LB1(f) peptide (from the P5-like fimbrin protein) were introduced followed by the 10 LB1(f) DNA sequence, followed by the DNA coding for the 5 natural residues C-terminal to the LB1(f) peptide. The plasmid (called LPD-LB1-A) is shown in Figure 3 and was made as follows:

pRIT 14588 was cleaved with *NcoI* and *SpeI*, and the linear large fragment was dephosphorylated. The LB1(f) peptide gene was amplified up from the ntHi-1128 P5-like 15 fimbrin gene with the following primers:

Primer LB-Baka-01 (5' - containing an *NcoI* site)

5'-CTA-GCC-ATG-GAT-GGT-GGC-AAA-GCA-GGT-G-3'

Primer LB-Baka-05 (3' - containing an *SpeI* site)

5'-CAC-TAG-TAC-GTG-CGT-TGT-GAC-GAC-3'

20

The DNA produced by PCR amplification was cleaved with *NcoI* and *SpeI*. The LB1(f) DNA fragment was purified, and ligated into the *NcoI* and *SpeI* sites of the cleaved pRIT14588. The ligation mixture was transformed into *E. coli* AR58, and the transformation product was spread onto solid medium (BP) LBT + Kanamycin (50  $\mu$ g/mL). The plates were incubated at 30°C overnight. Transformants were checked by 25 PCR, and positive candidates were grown in liquid culture at 30°C. In order to initiate expression of the LPD-LB1(f) chimeric polypeptide, the culture was subjected to a change in temperature from 30°C to 39°C during 4 hours. Expression was monitored on

a 12.5 % acrylamide gel (viewed either with Coomassie stain and/or Western Blot). The molecular size of the chimeric polypeptide was about 44 kDa.

Example 2b) Producing a LPD - LB1(f) Group 2 + LB1(f) Group 1 fusion

5 The plasmid (called LPD-LB1-II) is shown in Figure 4 and was made as follows:

Plasmid LPD-LB1-A was cleaved with *Nco*I and the linear DNA was dephosphorylated. The Group 2 LB1(f) peptide gene was amplified up from the ntHi-1715MEE P5-like fimbrin gene with the following primers:

Primer NT1715-11NCO (5' containing an *Nco*I site)

10 5'-CAT-GCC-ATG-GAT-GGC-GGT-AAA-GCA-GGT-GTT-GCT-3'

Primer NT1715-12NCO (3' containing an *Nco*I site)

5'-CAT-GCC-ATG-GCA-CGT-GCT-CTG-TGA-TG-3'

15 The DNA produced by PCR amplification was cleaved with *Nco*I. The LB1(f) DNA fragment was purified, and ligated into the open *Nco*I site of the cleaved LPD-LB1-A plasmid (5' to the gene for the Group 1 LB1(f) peptide). The ligation mixture was transformed into *E. coli* AR58, and the transformation product was spread onto solid medium (BP) LBT + Kanamycin (50 µg/mL). The plates were incubated at 30°C overnight. Transformants were checked by PCR, and positive candidates were grown in 20 liquid culture at 30°C. In order to initiate expression of the LPD-LB1(f)<sub>2,1</sub> chimeric polypeptide, the culture was subjected to a change in temperature from 30°C to 39°C during 4 hours. Expression was monitored on a 12.5 % acrylamide gel (viewed either with Coomassie stain and/or Western Blot). The molecular size of the chimeric polypeptide was about 50 kDa.

25

Example 2c) Producing a Lipoprotein D - LB1(f) Group 2 + LB1(f) Group 1 + LB1(f) Group 3 fusion

The plasmid (called LPD-LB1-III) is shown in Figure 5 and was made as follows:

Plasmid LPD-LB1-II was cleaved with *SpeI* and the linear DNA was dephosphorylated. The Group 3 LB1(f) peptide gene from ntHi-1929MEE was made by hybridising the following primers:

Primer NT1729-18 SPE (5' - containing a cleaved *SpeI* site at 5' end)

5 5'-CTA-GTC-GTT-CTG-ACT-ATA-AAT-TCT-ACG-ATA-ATA-AAC-GCA-TCG-  
ATA-GTA-3'

Primer NT1729-19 SPE (3' - containing a cleaved *SpeI* site at 3' end)

5'-CTA-GTA-CTA-TCG-ATG-CGT-TTA-TCG-TAG-AAT-TTA-TAG-GCA-GAA-  
CGA 3'

10

The hybridised DNA contained the gene for the Group 3 LB1(f) peptide and a cleaved *SpeI* at either end. The LB1(f) DNA fragment was ligated into the open *SpeI* site of the cleaved LPD-LB1-II plasmid (3' to the gene for the Group 1 LB1(f) peptide). The ligation mixture was transformed into *E. coli* AR58, and the transformation product was spread onto solid medium (BP) LBT + Kanamycin (50 µg/mL). The plates were incubated at 30°C overnight. Transformants were checked by PCR, and positive candidates were grown in liquid culture at 30°C. In order to initiate expression of the LPD-LB1(f)<sub>2,1,3</sub> chimeric polypeptide, the culture was subjected to a change in temperature from 30°C to 39°C during 4 hours. Expression was monitored on a 12.5 % acrylamide gel (viewed either with Coomassie stain and/or Western Blot). The molecular size of the chimeric polypeptide was about 53 kDa.

20  
25 Example 2d) Characterisation of the Expression of the Chimeric Polypeptides  
Expression of the above chimeric polypeptides was monitored on a 12.5 % acrylamide gel which was observed as either:

- a) a Coomassie stained gel (Figure 6)
- b) a Western blot
  - 1) using rabbit anti-LB1 (Figure 7)
  - 2) using a Mab anti-LPD (Figure 8)

3) using an Ab against the six-Histidine Purification Tag (Figure 9)

As can be observed, each chimeric polypeptide can be expressed efficiently in *E. coli*.

5

#### Example 3: Purification of the Chimeric Polypeptides

The purification of LPD-LB1(f)<sub>2,1,3</sub> (expressed using the construct shown in Figure 5) was achieved as follows.

The *E. coli* were washed and resuspended in phosphate buffer (50 mM, pH 7.0).

10 The cells were lysed by gently swirling them overnight at 4 °C in the presence of 3% Empigen. The solution was then centrifuged for 30 minutes at 8,000 rpm in a Beckman JA10 rotor. The supernatant was diluted 4 times in 50 mM phosphate buffer, 500 mM NaCl, pH 7.0. The first stage of purification was achieved on a Qiagen NTA Ni<sup>++</sup> column due to the presence of the six histidine tag at the C-terminus of the polypeptide.

15 The column was equilibrated with 10 mM sodium phosphate buffer, 500 mM NaCl, 0.5% Empigen, pH 7.5, and the polypeptide was eluted off the column with an imidazole gradient (0-100 mM) in 20 mM sodium phosphate buffer, 0.5% Empigen, pH 7.0. Elution was followed by running fractions on SDS-PAGE gels.

The next step in the purification was on a Bio-Rad Macro-Prep 50S column. The 20 polypeptide bound to the column equilibrated in 20 mM phosphate buffer, 0.5 % Empigen, pH 7.0, and was eluted from the column using a gradient of 0 to 500 mM NaCl in the same buffer. Elution was followed by running fractions on SDS-PAGE gels.

The last (polishing) step of the process was done using a Sephadex S200 HR size 25 exclusion column. The polypeptide solution from the previous step was firstly concentrated with a Filtron Omega 10 kDa concentrator device. The resulting solution was loaded and run on the column equilibrated with PBS buffer with 0.5% Empigen. Elution of the polypeptide was followed by running fractions on SDS-PAGE gels.

The pooled fractions were filtered through a 0.22 µm filter. The resulting protein runs as one pure band on a Coomassie stained SDS-PAGE gel, and the equivalent

Western blot using an anti-LB1 antibody. Tests showed that the protein remained intact even after 7 days at 37 °C.

Approximately 200 mg of polypeptide per litre of cell culture can be purified by this method.

5

Example 4: Preclinical Experimentation on vaccine effectiveness of the chimeric polypeptides

Example 4a) Generation of antisera

Antisera was generated against 4 types of antigen: LPD; PD; LPD-LB1(f)<sub>2,1,3</sub> 10 (made recombinantly using plasmid LPD-LB1-III); LB1 (a group 1 LB1(f) peptide fused to a T-cell promiscuous epitope from measles virus fusion protein, the sequence of the peptide being: RSDYKFYEDANGTRDHKKGPSLKLLSLIKGVIVHRLEGVE).

Four cohorts comprising 5 chinchillas were immunised, each cohort with one of the immunogens identified above. The dosage was 10 µg antigen / 200 µL AlPO<sub>4</sub> / 20 µg 15 MPL (3-O-deacylated monophosphoryl lipid A) for the first three antigens, and 10 µg antigen delivered in Complete or Incomplete Freund's Adjuvant (CFA or IFA) for LB1.

A total of three doses were injected at one month intervals.

Example 4b) Passive Transfer and Challenge

20 This study aimed to perform an *in vivo* challenge study of passively immunised chinchillas to determine the relative efficiency among the 4 immunogen (or sham) formulations to facilitate clearance of ntHi from the nasopharynx.

Five cohorts of 11 chinchillas each (*Chinchilla laniger*) free of middle ear disease were inoculated intranasally on day -7 with 6 x 10<sup>6</sup> TCID<sub>50</sub> adenovirus type 1. On day -1 25 each chinchilla cohort was passively immunised with a 1:5 dilution of one of the four serum samples described in Example 4a via cardiac puncture. The fifth cohort (the sham) received pyrogen-free sterile saline solution by cardiac puncture instead. About 5 mL serum / kg animal was administered.

On day 0 the cohorts were intranasally challenged with ntHi: about  $10^8$  cfu ntHi #86-028NP (group 1) per animal. Statistical evaluation of the passive transfer study was performed prior to de-blinding the study.

This sequential inoculation with two pathogens closely mimics both the natural 5 route of acquisition of these agents as well as their synergistic interaction in the human host.

The severity of the disease was scored by otoscopic observation. This is rated on a 0-4 scale. Signs of tympanic membrane (TM) inflammation were observed to obtain a score: the presence of effusion, small vessel dilation, air-fluid interface, opacity, etc.

10 A repeated measures analysis of variance was used to compare the pattern of responses over time (days) and ear (left or right) for the five groups (cohorts). Due to the large number of repeat observations on each animal, the analysis was divided into 5 sections as follows: days 1-7, days 8-14, days 19-21, days 22-28, and days 29-33. There was little variation in the responses on days -7 through 0, 34 and 35 and therefore no 15 such analysis was performed on those times. Where possible (when there was non-zero variability in the mean response), tests were performed to compare the mean responses between the groups at these time points. Tukey's HSD test was used for all post-hoc multiple comparisons. Significance was assessed using an alpha level of 0.05.

20 The results are shown in Figure 10. Inflammation increases over time for all groups in a significant manner during the period of day 1 to 7. During days 29-33 inflammation decreased over time in a significant manner for all groups. As can be seen from the data, the serum containing antibodies against recombinant LPD-LB1(f)<sub>2,1,3</sub> helped to reduce the TM inflammation throughout the experiment. An effective 25 vaccinogen should maintain TM inflammation at or below 1.5 for the duration of the study period. LPD-LB1(f)<sub>2,1,3</sub> anti-serum only allowed the mean inflammation score to rise above 1.5 for 2 days as well as inducing a consistent downward trend thereafter.

In addition to otoscopy, tympanometry (which measures changes in middle ear pressure) was also employed. These two measurements can be used in conjunction to give a reliable indication of whether an effusion has taken place in a middle ear. Figure

11 shows the results of this analysis. Clearly, the recombinant LPD-LB1(f)<sub>2,1,3</sub> performed well in this study when considering the outcome measures of preventing both TM inflammation and the development of effusion. Overall LPD-LB1(f)<sub>2,1,3</sub> ranks second only to the positive control, the LB1 peptide. The LB1 peptide, however, was adjuvanted 5 with CFA (a very strong adjuvant) and can therefore not be directly compared to the LPD-LB1(f)<sub>2,1,3</sub> result.

A statistical evaluation on the data presented in Figure 11 is presented in Table 10. The evaluation compared the reduction in percent effusion in each immunized cohort to that observed in sham immunized animals during peak incidence of disease [the four 10 days of observation in which at least 50% of sham ears contained an effusion (had otitis media)].

The positive control (anti-LB1/CFA) was significant at  $p < 0.001$  on all four days (days 11 - 14). Anti-LPD-LB1(f)<sub>2,1,3</sub> inhibited the development of otitis media at a p-value < or equal to 0.001 on days 11, 12, 13 and 14 also. Anti-PD was significant on 15 days 13 and 14 only whereas anti-LPD was able to prevent the development of otitis media relative to sham animals on day 14 only ( $p$  value close to 0.02).

The recombinant LPD-LB1(f)<sub>2,1,3</sub> polypeptide therefore significantly inhibits the development of otitis media in chinchillas which were passively transferred with this serum pool.

Day	Group	% Effusion	p-value
11 (Sham = 70%)	LB1	0	<0.0001
	PD	45	0.1010
	LPD-LB1(f)213	17	0.0010
	LPD	68	0.8886
12 (Sham = 80%)	LB1	0	<0.0001
	PD	55	0.0854
	LPD-LB1(f)213	22	0.0004
	LPD	68	0.3788
13 (Sham = 65%)	LB1	15	0.0012
	PD	18	0.0020
	LPD-LB1(f)213	17	0.0002
	LPD	41	0.1188
14 (Sham = 60%)	LB1	0	<0.0001
	PD	5	0.0002
	LPD-LB1(f)213	0	<0.0001
	LPD	23	0.0146

**Table: 10** - A comparison of % ears containing effusion in the LB1, PD, LPD-LB1(f)213, and LPD groups with % of ears containing effusion in the Sham group on days 11 through 14..

#### Example 4c) Adherence Inhibition data

An established single cell adherence assay was carried out using human oropharyngeal cells. The mean percent inhibition of adherence ( $\pm$ sem) of ntHi strains to these cells by the immune chinchilla sera produced in Example 4a. The results using anti-sera against LPD-LB1(f)<sub>2,1,3</sub> and LPD can be seen in Table 11. The anti-sera against LPD-LB1(f)<sub>2,1,3</sub> was seen to be effective at inhibiting adherence of Group 1 and Group 2 ntHi strains. It was also more effective against all the strains than anti-LPD serum was.

Cohort Name	ntHi strain (Group)	n	Pooled Serum Dilution					
			1:25	1:50	1:100	1:200	1:400	1:800
LPD/ AlPO <sub>4</sub> / MPL	86-028L (Group 1)	3	29±3	31±4	13±7	19±8	12±5	16±7
	1128MEE (Group 1)	2	0±0	12±12	8±5	12±1	8±8	16±1
	266NP (Group 2a)	3	46±9	38±7	24±13	24±21	30±16	28±19
LPD- LB1(f) <sub>213</sub> / AlPO <sub>4</sub> / MPL	86-028L (Group 1)	3	32±2	36±1	38±2	27±3	3±2	19±3
	1128MEE (Group 1)	2	24±14	23±4	30±7	13±13	11±11	12±6
	266NP (Group 2a)	3	52±10	43±3	36±7	13±10	6±9	14±19

**Table: 11** - The mean percent inhibition of adherence ( $\pm$ sem) of ntHi strains to human oropharyngeal cells by immune chinchilla sera.

5

Although certain embodiments of this invention have been shown and described, various adaptations and modifications can be made without departing from the scope of the invention as described in the appended claims. For example, peptides or polypeptides having the substantially the same amino acid sequence as described herein are within the scope of the invention.

SEQ ID NO: 1

**RSDYKFYEAANGTRDHKKG**

[from strain ntHi-10567RM (Group 1 type)]

5 SEQ ID NO: 2

**RSDYKLYNKNSSNSTLKNLGE**

[from strain ntHi-1715MEE (Group 2a type)]

SEQ ID NO: 3

10 **RSDYKFYDNKRID**

[from strain ntHi-1729MEE (Group 3 type)]

SEQ ID NO: 4

**RSDYKLYNKNSSTLKDLGE**

15 [from strain ntHi-183NP (Group 2b type)]

SEQ ID NO: 5

**RSDYKFYEDANGTRDHKKG**

[from strain ntHi-1128 (Group 1 type)]

20

SEQ ID NO: 6

**RSDYKFYEAAPNSTRDXKKG**

[from protein P5 from ntHi residues 119-137 (Group 1 type)]

We claim:

1. A peptide corresponding to LB1(f) peptides from the P5-like fimbrial protein of *Haemophilus influenzae* comprising one or more amino-acid sequences selected from the 5 group consisting of:
  - SEQ. ID NO. 1,
  - SEQ. ID NO. 2,
  - SEQ. ID NO. 3, and
  - SEQ. ID NO. 4
- 10 or any antigenically related variants of said sequences which have an identity of at least 75% and are capable of immunologically mimicking the corresponding antigenic determinant site of the P5-like fimbrial protein of non-typeable *Haemophilus influenzae*, with the proviso that the antigenically related variants does not include those peptides provided in SEQ ID NO:5 or SEQ ID NO:6.
- 15 2. A peptide of claim 1 which comprises the amino-acid sequence provided in SEQ ID NO:1.
3. A peptide of claim 1 which comprises the amino-acid sequence provided in SEQ ID 20 NO:2.
4. A peptide of claim 1 which comprises the amino-acid sequence provided in SEQ ID NO:3.
- 25 5. A peptide of claim 1 which comprises the amino-acid sequence provided in SEQ ID NO:4.
6. A chimeric polypeptide comprising one or more peptides of claims 1-5 conjugated to a carrier polypeptide which comprises at least one T-cell epitope.

7. A chimeric polypeptide of claim 6 which also comprises a purification tag peptide sequence.
- 5 8. A chimeric polypeptide of claim 7 wherein the purification tag peptide sequence is a Histidine-tag sequence.
9. A chimeric polypeptide of claim 6 wherein the carrier polypeptide is lipoprotein D.
- 10 10. A chimeric polypeptide of claim 6 wherein the amino acid sequences of the LB1(f) peptide subunits used are selected from the group consisting of SEQ ID NO:1, 2, or 3.
11. A chimeric polypeptide comprising three LB1(f) subunits and lipoprotein D, wherein the amino acid sequences of the LB1(f) subunits used are provided in SEQ ID NO: 2, 3  
15 and 5.
12. A chimeric polypeptide of claim 11 which also comprises a Histidine purification tag sequence.
- 20 13. A chimeric polypeptide of claim 11 wherein the order of the peptide components from the N-terminus of the polypeptide is: lipoprotein D, LB1(f) subunit (SEQ ID NO: 2), LB1(f) subunit (SEQ ID NO: 5), and LB1(f) subunit (SEQ ID NO: 3).
14. A vaccine composition comprising an immunogenic amount of at least one peptide or  
25 polypeptide from claims 1-13 in a pharmaceutically acceptable excipient, and an optional adjuvant.

15. The use of an immunogenic amount of at least one peptide or polypeptide from claims 1-13 in a pharmaceutically acceptable excipient, and an optional adjuvant, to prevent or treat *Haemophilus influenzae* disease.
- 5 16. The use of claim 15 wherein the *Haemophilus influenzae* disease is *otitis media*, sinusitis, conjunctivitis, or lower respiratory tract infection.
- 10 17. A method of inducing immunity in a mammal susceptible to *Haemophilus influenzae* infection comprising the administration to the mammal of an effective amount of the vaccine according to claim 14.
18. A method of preventing *Haemophilus influenzae* infection comprising the administration to a mammal an effective amount of a vaccine according to claim 14.
- 15 19. A DNA or RNA molecule encoding one of the LB1(f) peptides or polypeptides provided in claims 1-13.
- 20 20. A DNA or RNA molecule of claim 19 wherein the DNA sequence of said LB1(f) polypeptide is provided in Figure 5.
21. A DNA or RNA molecule of claim 19 contained within an expression vector, wherein said expression vector is capable of producing said LB1(f) peptide or polypeptide when present in a compatible host cell.
- 25 22. A DNA or RNA molecule of claim 20 contained within an expression vector, wherein said expression vector is capable of producing said LB1(f) polypeptide when present in a compatible host cell.
23. A host cell comprising the expression vector of claim 21 or 22.

24. A process for producing a LB1(f) peptide or polypeptide comprising culturing the host cell of claim 23 under conditions sufficient for the production of said polypeptide and recovering the LB1(f) peptide or polypeptide.

5

25. A process for producing a LB1(f) peptide or polypeptide of claim 24 wherein the process comprises the steps of lysing the host cells, and purifying the soluble extract using an immobilised Nickel column step, a cation exchange column step, and a size exclusion column step.

10

26. A process for producing a host cell which produces a LB1(f) peptide or polypeptide thereof comprising transforming or transfecting a host cell with the expression vector of claim 21 or 22 such that the host cell, under appropriate culture conditions, expresses a LB1(f) peptide or polypeptide.

15

27. A purified antibody which is immunospecific to a LB1(f) peptide provided in claims 1-5.

20

28. A method of detecting the presence of *Haemophilus influenzae* in a sample by contacting said sample with the antibody of claim 27 in the presence of an indicator.

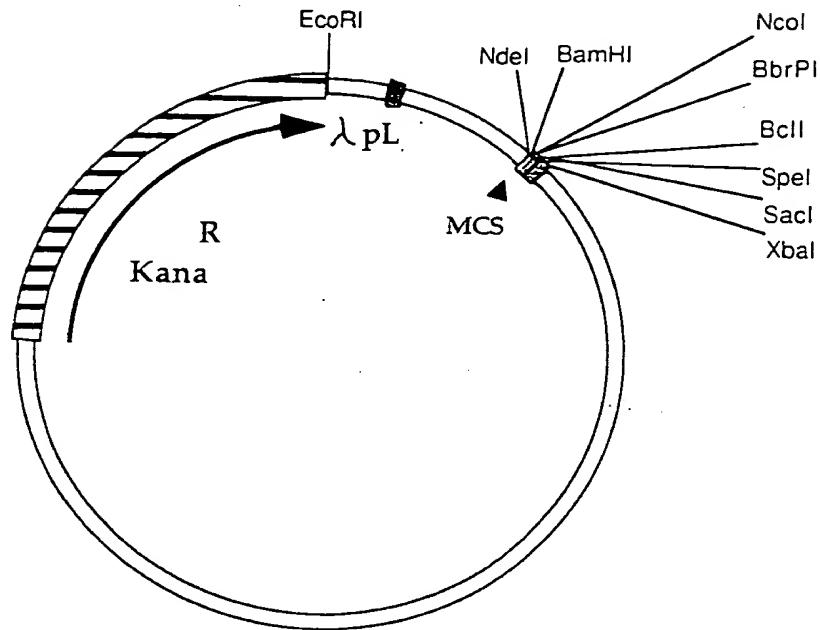
25

29. A method of detecting the presence of *Haemophilus influenzae* in a sample by contacting said sample with a DNA probe or primer constructed to correspond to the wild-type nucleic acid sequence which codes for a LB1(f) peptide of the P5-like fimbrial protein of *Haemophilus influenzae*, characterised in that the probe is selected from the group consisting of gene sequences as provided in Tables 6-8.

30. A reagent kit for diagnosing infection with *Haemophilus influenzae* in a mammal comprising the DNA probes of claim 29 or a LB1(f) peptide of claims 1-5 or an antibody of claim 27.

THIS PAGE BLANK (USPTO)

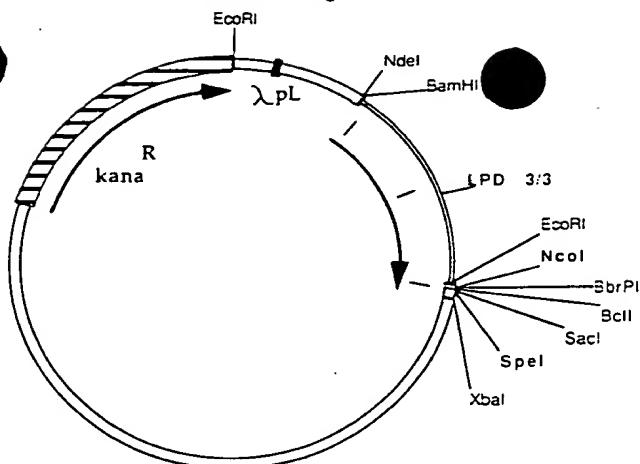
pMGMCS (figure : 1)



CTCTTACACATTCCAGCCCTGAAAAAGGGATCAATTAAACCACACCTTARGGAGGATATACATA  
► M  
NdeI  
BamHI NcoI BbrPI BclI SacI SphI  
TGGATCCCATGGCCACGTGTGATCAGAGCTCAACTAGTGGCCACCATCACCACCA  
► et Asp Pro Met Ala Thr Cys Asp Glu Ser Ser Thr Ser Gly His His His His His  
XbaI  
TTAATCTAGAATCGATAAGCTTCGACCGATGCC  
► s... .

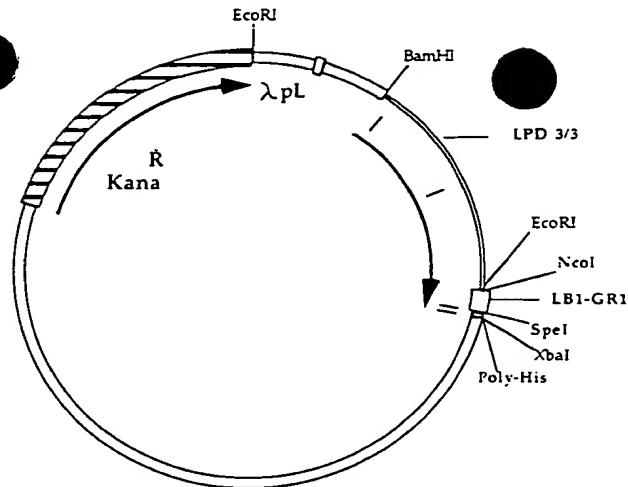
THIS PAGE BLANK (USPTO)

pRIT 14588 (figure : 2)



Ndel BamHI  
 CTCTTACACATTCAGCCCTGAAAAGGGCATCAAATTAAACCACACCTTAAGGAGGATAAACATATCGAT  
 ▶ Met Asp  
 CCAAAACTTCTGCTTCTTATTAGCAGCTGGCGTACTACACGGTTAGCAGCCATTCACTAAATATG  
 ▶ ProLysThr LeuAl aLeuSer LeuLeuAl aAl aGlyVal LeuAl aGlyCysSer Ser HisSer AsnMet  
 GCGAATACCCAAATGAAATCAGACAAATCATTATTGCTCACCGTGGCTACCGGTTATTACAGAGGAT  
 ▶ Al aAsnThr GluMet LysSer AspLysIleIleAlaHisArgGlyAlaSer GlyTyrLeuProGluHis  
 ACGTTAGAACCTAAAGCACTGGCTTGACACAGGGTGATTATTAGAGCAAGATTACCACTGACTAAG  
 ▶ Thr LeuGluSer LysAl aLeuAl aPheAl aGluAl aAspTyrLeuGluGlu nAspLeuAl aMet Thr Lys  
 GATGGTCGTTAGTGGTTATTCAAGCATCTTTAGATGGCTGACTGATCTGGGAATTTCACCATGACTAAG  
 ▶ AspGlyArgLeuValValIleHisAspHisPheLeuAspGlyLeuThrAspValAlaLysLysPheProHis  
 CGTCATCGAAAGATGGCGTTACTATGTCATCGACTTTACCTAAAGAAATTCAAAGTTAGAAATGACA  
 ▶ ArgHisArgLysAspGlyArgTyrTyrValIleAspPheThrLeuLysGluValLeGluSer LeuGluMet Thr  
 GAAACTTTGAAACCAAGATGGCAACAGCCCAAGTTATCCTAATGGTTCCCTTTGGAAATCACAT  
 ▶ GluAsnPheGluThr LysAspGlyLysGluAlaGluValTyrProAsnArgPheProLeuTrpLysSer His  
 TTAGATTCAACCTTTGAGATGAAATTGAAATTCAAGCTTAGAATTCCTGGAAATTCACTGGAAATTCA  
 ▶ PheArgIleHisThr PheGluAspGluIleGluPheIleGluGlyLeuGluLysSer Thr GlyLysLysVal  
 CGGATTATCCAGAAATCAAAGCACCTGGTCCACCATCAAATGGTAAAGATATTGCTGGAAACCTC  
 ▶ GlyIleTyrProGluValLysAlaProTrpPheHisHisGluAspIleAlaAlaGluThrLeu  
 LPD 3/3  
 AAAGTGTAAAAAAATATGGCTATGATAAGAAAACGATATGGTTACTTACAAACTTTCGATTTATGAA  
 ▶ LysValLeuLysLysTyrGlyTyrAspLysThrAspMetValTyrLeuGluThrPheAspPheAsnGlu  
 TTAAACGTATCAAACCGAATTACTCCACAAATGGGATGGATTGAAATTAGTCATAATTGGCTTAT  
 ▶ LeuLysArgIleLysThr GlyLeuLeuProGluMetGlyMetAspLeuLysLeuValGluLeuIleAlaTyr  
 ACAGATTGGAAAGGAAACAAAGAAAAAGACCCAAAGGGTTATGGTAAACTATAATTGATGGATGTT  
 ▶ ThrAspTrpLysGluThrGluGluLysAspProLysGlyTyrTrpValAsnTyrAsnTyrAspTrpMetPhe  
 AACCTGGTGCATGGCAGAAGTGGTAAATATCCCGTGGTGGCCAGGGTATATGTTAGTTAT  
 ▶ LysProGlyAlaMetAlaGluValValLysTyrAlaAspGlyValGlyProGlyTyrTrpMetLeuValAsn  
 AAAAGAAGATCCAAACCTGATAATATGTGACACTCCGGTGGTAAAGAATTCGCACAAATAATGTGAA  
 ▶ LysGluGluSerLysProAspAsnIleValTyrThrProLeuValLysGluLeuAlaGluTyrAsnValGlu  
 GTGCATCCTTACACCGTGGTAAGATGCTGACCTGCCGAGTTTACAGACGCTAAATGAAATTGTATGATGCC  
 ▶ ValHisProTyrThrValIleLysAspAlaLeuProGluPhePheThrAspValAsnGluMetTyrAspAla  
 EcoRI  
 TTATTGAAATAATCAGGGCCACAGGGTGTATTACTGTTCCAGATACTGGCGTGGAAATTCTTAAAGGA  
 ▶ LeuLeuAsnLysSerGlyAlaThrGlyValPheThrAspPheProAspThrGlyValGluPheLeuLysGly  
 Ncol BbrPI BclI SacI Spel  
 ATAAATTCATGCGCACCGTGTGATCAGAGCTCAACTAGTGGCCACCATCACCATCACCATTA  
 ▶ IleLysSerMetAlaThrCysAspGluSerSerThrSerGlyHisHisHisHisHis...  
 XbaI  
 TCTAGAATCGATAAGCTCGACCGATGCC

**THIS PAGE BLANK (USPTO)**



BamHI  
 CTCTTACACATTCCAGCCCTGAAAAAGGGCATCAAATTACCCACACCTTAAGGAGGATAAACATHTGGATCCAA  
 AACTTACCCCTTTCTTATTACCAAGCTGGCTACTACCCAGGTTGAGCAGCCATTCACTAAATATGGCGAATACCC  
 ▶ sThr LeuAl aLeuSer LeuLeuAl aAl aGlyVal LeuAl aGlyCysSer Ser HisSer AsnMetAl aAsnThr G  
 AATGAAATCAGACAAATCATTATTGCTACCGTGGCTAGCGGTATTTACCGAGGATAAGCTAGATCTAA  
 ▶ I nMetLysSerAspLysIleIleIleAlaHisArgGlyAlaSerGlyTyrLeuProGlyHi sThr LeuGlySer Lys  
 GCACTTGGCTTGGCACACAGGCTGATTATTTAGAGCAAGATTTAGCAATGACTAAAGGATGGTGGTTAGTGGTAT  
 ▶ Al aLeuAl aPheAl aGlyAl nAl aAspTyrLeuGlyAl nAspLeuAl aMet Thr LysAspGlyA rgLeuVal Val II  
 TCACGATCACCTTTAGATGGCTTGACTGATGGCGAAATTCACATCGTAAAGATGGCGTTACT  
 ▶ eHi sAspHi sPheLeuAspGlyLeuThrAspValAlaLysPheProHi sArgHi sArgLysAspGlyA rgTyrT  
 ATGTCATCGACTTACCTTAAAGAAATTCAACTTACGAAATGACAGAACTTGAAGATGAAATTGAA  
 ▶ yVal IleAspPheThrLeuLysGlyIleGlySer LeuGlyMet Thr GlysAsnPheGlyThr LysAspGlyLysGly n  
 GCGCAAGTTATCCTATCGTTCCCTCTTGGAAATCACATTAGAAATTGACATACCTTGAAGATGAAATTGAA  
 ▶ Al aGlyVal TyrProAsnArgPheProLeuTrpLysSer HisPheArgIleHisThr PheGlyAspGlyIleGlyPh  
 TATCCGAGCTTGGAAATTCGACTGGAAATGGAGATTATCCGAAATTCGACACCTTGGTGGCGACATC  
 ▶ eIleGlyLeuGlyLeuLysSer Thr GlysLysVal GlysIleTyrProGlyIleLysAlaProTrpPheHisGly  
 LPD 3/3  
 AAAATGCTAAAGATMTGCTGCTGAAACGCTAAAGCTTAAAAAAATATGGCTATGATAAACAAACGGATATGCTT  
 ▶ I nAsnGlyLysAspIleAlaAlaGlyThr LeuLysVal LeuLysLysTyrGlyTyrAspLysLysThrAspMetVal  
 TACTTACAAAATTCGATTGAAATTGAAATACGGTACAAACGGAAATTACTTCCACAAATGGAAATGGATTGAA  
 ▶ TyrLeuGlyThr PheAspPheAsnGlyLeuLysArgIleLysThr GlysLeuLeuProGlyMetGlyMetAspLeuLys  
 ATTAGTTCAATTAAATGGCTTACAGATTGGAAAGAACACAGAAAGACCCAAAGGTTATTGGTAAACTATA  
 ▶ sLeuValGlyIleAlaTyrThrAspTrpLysGlyThrGlyLysAspProLysGlyTyrTrpValAsnTyrA  
 ATTACGATTGGATGTTAAACCTGGCAATGGCAGAAGTGGTAAATATGCCATGGTGTGGCCACAGGTTGGTAT  
 ▶ snTyrAspTrpMetPheLysProGlyAlaMetAlaGlyValValIlysTyrAlaAspGlyValGlyProGlyTrpTyr  
 ATGTTAGTTAATAAGAAGAATCCTAACCTGATAATTGTTGACACTCCGTTGGTAAAGAACTTGGACAAATATAAA  
 ▶ MetLeuValAsnLysGlyGlySerLysProAspAsnIleValTyrThrProLeuValLysGlyLeuAlaGlynTyrAs  
 TGTGGAAATGTCATCCCTAACCGTGGCTTAAAGTGCATGCCGAGTTTTGACAGACGTAATCAATGTATGATG  
 ▶ nValGlyValHisProTyrThrValArgLysAspAlaLeuProGlyPhePheThrAspValAsnGlyMetTyrAspA  
 EcoRI  
 CCTTATTGAAATAATCAGGGCAACAGGTGTTTACTGATTTCCACGATACTGGCTGGAAATTCTAAAGGAATA  
 ▶ I aLeuLeuAsnLysSerGlyAlaThrGlyValPheThrAspPheProAspThrGlyValGlyPheLeuLysGlyIle  
 Ncol  
 AAATCAGCGATGGGGCAAGCGGTTGGCTTACTGATCTGATTATAAAATTGAAAGATGCAA  
 ▶ LysSerMetAspGlyGlyLysAlaGlyValAlaLeuValArgSerAspTyrLysPheTyrGlyAspAlaAs  
 SpeI Poly-His  
 CGGTACTCGTGACCACAAAGAAAGGTGGTCAACACACGACTAGTGGCCACCATCACCATCACCATT  
 ▶ nGlyThrArgAspHisLysGlyArgHisThrAlaArgThrSerGlyHisHisHisHis  
 XbaI  
 AATCTAGAATCGATAAGCTTCGACCGATGCC  
 ||  
 ▶ ...

THIS PAGE BLANK (USPTO)

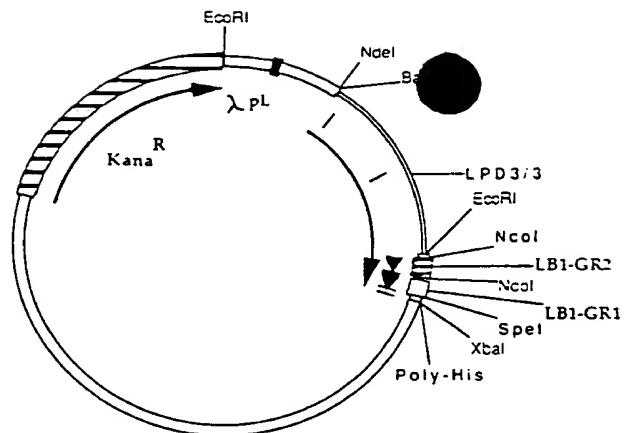
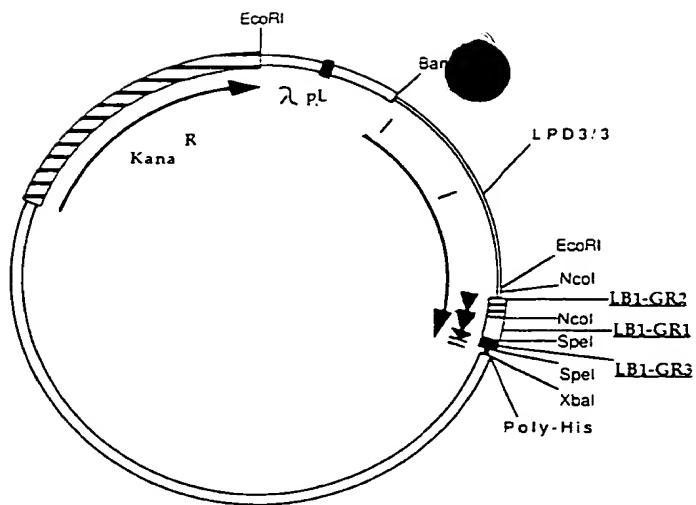


fig 4

NdeI BamHI  
 CTCTTACACATTCGAGCCCTGAAAAAGGGCATCAATTAAACCAACCTTAAGGAGGTATAACATATGGATCCAAAAACT  
 ↓ MetAspProLysThr  
 TTAGCCCTTTCTTATTAGCAGCTGGCGTACTACCAGGTTGAGCACCCATTCACTCAAAATAGCGAATACCCAAATTAAA  
 ▶ LeuAlaLeuSer LeuLeuAlaAlaGlyValLeuAlaGlyCysSer Ser HisSer SerAsnMetAlaAsnThr GluMetLys  
 TCAGACAAATCATTTATGCTCACCGTGGCTACCGGTTATTACCGAGGATACCTAGAATCTAAACGACTTGGCTTT  
 ▶ SerAspLysIleIleAlaAlaHisArgGlyAlaSerGlyTyrLeuProGluHiSerLysAlaLeuAlaPhe  
 GCACGACAGGCTGATTATTAGACCAAGATTACCAAGTACTAAGGATGGTGGTTAGTGGTTATACGATTACTTTTA  
 ▶ AlaGluGlnAlaAspTyrLeuGluGlnAspLeuAlaMetThrLysAspGlyArgLeuValValIleHisAspHisPheLeu  
 GATGGCTTGACTGATGTTGCGAAAAATTCCACATCGTAAAGATGGCGTTACTATGTCATCGACTTACTTA  
 ▶ AspGlyLeuThrAspValAlaLysLysPheProHisArgHisArgLysAspGlyArgTyrTyrValIleAspPheThrLeu  
 AAAGAAATTCAAAAGTTAGAATATCAGAAAACCTTGAACACGATGGCAAACAGGCGAAAGTTTATCTTAATGGTT  
 ▶ LysGluIleGluSerLeuGluMetThrGluAsnPheGluThrLysAspGlyLysGlnAlaGlnValTyrProAsnArgPhe  
 CCTCTTGGAAATCACATTAGATTGATACCTTGAGATGAAATTGAAATTGAAAGGCTAGAATATCCACTGGC  
 ▶ ProLeuTrpLysSerHisPheArgIleHisThrPheGluAspGluIleGluPheIleGluGlyLeuGluLysSerThrGly  
 AAAAAGTAGGGATTATCCAGAAATCAGAACCTTGCTGACCATCAAAATGCTAAAGATATTGCTCTGAAACGTC  
 ▶ LysLysValGlyIleTyrProGluIleLysAlaProTrpPheHisGlyLysAspIleAlaAlaGluThrLeu  
 LPD 3/3  
 AAAGTGTAAAAAAATATGCGTATGATAAGAAAAACCGATATGGTTACTTACAAATTTCGATTTAAATGAAATTAACG  
 ▶ LysValLeuLysLysTyrGlyTyrAspLysLysThrAspMetValTyrLeuGluThrProAspProAsnGluLeuLysArg  
 ATCRAACGGAACTACTCCACAAATGGGAATGGATTGAAATTACTCAATTGCTATAACGATTGCAAAGAACACA  
 ▶ IleLysThrGluLeuLeuProGlnMetGlyMetAspLeuLysLeuValGluLeuIleAlaTyrThrAspTrpLysGlyThr  
 CAAGAAGAACCCAAAGGGTTATCCGTTAACTATAATTACGATTGGATTTAAACCTGCTATAACGAGAACGTT  
 ▶ GluGlnLysAspProLysGlyTyrTrpValAsnTyrAsnTyrAspTrpMetPheLysProGlyAlaMetAlaGluValVal  
 AAATATGCCGATGGTGGCCAGGTTGCTATGTTAGTTAAAGAAGAATCCAAACCTGATAATTGTGTTACACT  
 ▶ LysTyrAlaAspGlyValGlyProGlyTrpTyrMetLeuValAsnLysGluGluSerLysProAspAsnIleValTyrThr  
 CCGTGGTAAAGAAGACTTGCAAAATATAATGTGGAAGTCATCCTAACCCGCGTAAAGATGACTGCCGAGTTTTC  
 ▶ ProLeuValLysGluLeuAlaGlyTyrAsnValGluValHisProTyrThrValArgLysAspAlaLeuProGluPhePhe  
 ACAGACGTTAAATCAGATGATGCTTATGAAATCAGGGCAACAGGTTATTACTGATTTCCAGATACTGGC  
 ▶ ThrAspValAsnGlnMetTyrAspAlaLeuLeuAsnLysSerGlyAlaThrGlyValPheThrAspPheProAspThrGly  
 EcoRI Ncol LB1-GR2  
 CTGGAACTCTTAAATGGGATAAAATGCGATGGCTGAAACGCGTGTGGTTACTTCGTTGACTGAAATGGTAAATTGAC  
 ▶ ValGluPheLeuLysGlyIleLysSerMetAspGlyGlyLysAlaGlyValAlaLeuValArgSerAspTyrLysLeuTyr  
 Ncol  
 AACAAAAATAGTAGTACTAAAGTACTCTTAAACCTAGCCGAAACATGACAGGCGACSTGCCATGGTGGTCCGAAACCA  
 ▶ AsnLysAsnSerSerAsnSerThrLeuLysAsnLeuGlyGluHisArgAlaArgAlaMetAspGlyGlyLysAla  
 LB1-GR1  
 CGTGGTGGCTTGTGACCTTCGATTATAATTGAGATGCAACCGTACTCGTACCCACAGGAAAGCTGTCACACA  
 ▶ GlyValAlaLeuValArgSerAspTyrLysPheTyrGluAspAlaAsnGlyArgAspHisLysLysGlyArgHisThr  
 Spel Poly-His XbaI  
 GCACGTTACTGTCGGCAGCATGACCATGACCATTAATGAGATGATAAGCTTCGAAAGCATGCC  
 ▶ AlaArgThrSerGlyHisHiHisHiHisHiHis...

THIS PAGE BLANK (USPTO)

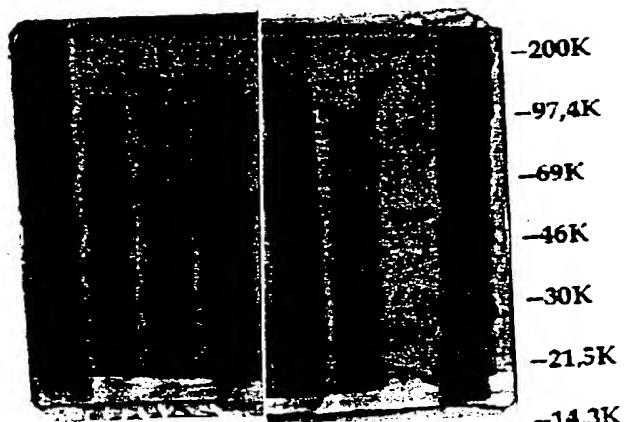


BamHI  
 CTCTTACACATTCAGCCCTGAAAAAGGGCATCAATTAAACCAACCTTAAGGAGGATNTACATATGGATCCA  
 ▶ Met Asp Pro Lys  
 ACTTTAGCCCTTCCTTATTAGCAAGCTACTAGCAGGTGTAGCAGGCCATCATCAATTATGCCAATACCCAA  
 ▶ Thr Leu Al aLeu Ser Leu Leu Al aAl aGly Val Leu Al aGly Cys Ser Ser His Ser Asn Met Al aAsn Thr Glu  
 ATGAAATCAGACAAAATCATTATTGCTCACCGTGTCTACCGGTTATTACAGAGCATACTGAGTAACTAAAGCA  
 ▶ Met Lys Ser Asp Lys Ile Ile Al aHi s Arg Glu Al aSer Glu Tyr Leu Pro Glu Hi s Thr Leu Glu Ser Lys Al a  
 CTTGCGTTGCACACAGGTGATTATTAGACAGATTACGATGACTAAGGATGGTGTCTTAGTGGTTATTAC  
 ▶ Leu Al aPhe Al aGln Glu Al aAsp Tyr Leu Glu Gln Asp Leu Al aMet Thr Lys Asp Glu Arg Leu Val Val Ile His  
 GATCACTTTTATAGTGGCTTGTACTGATGTTGCGAATTTCCACATCGTATCGTAAGATGGCCGTACTARGTC  
 ▶ Asp His Phe Leu Asp Glu Leu Thr Asp Val Al aLys Lys Phe Pro His Arg His Arg Lys Asp Glu Arg Tyr Val  
 ATCGACTTTACCTTAAAGAAATTCAAACCTTACGAAATGACAGAAACTTGAACACAAAGATGCCAAACAGGGAA  
 ▶ Ile Asp Phe Thr Leu Lys Glu Ile Glu Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Lys Asp Glu Lys Glu Al aGln  
 CCTTATCCTAATCGTTCCCTCTTGGAAATCAGATTAGAACTCATCTTGAAGATGAAATTGAATTATGAA  
 ▶ Val Ile Asp Pro Asn Arg Phe Pro Leu Trp Lys Ser His Phe Arg Ile His Thr Phe Glu Asp Glu Ile Glu Phe Ile Glu  
 GCCTTAGAAAATCCACTGGAAAAAAGTGGATTATCCAGAAATCAGACCTTGGTTCACCATCGAT  
 ▶ Glu Leu Glu Lys Ser Thr Glu Lys Lys Val Glu Ile Tyr Pro Glu Ile Lys Al aPro Trp Phe His Glu Asn Glu  
 AAAGATATTGCTGCTGAAACGCTAAAGTGTAAATATGGCTATGATAAGAAAACCGATAATGGTTACTTACAA  
 ▶ Lys Asp Ile Al aAl aGlu Thr Leu Lys Val Leu Lys Lys Tyr Glu Tyr Asp Lys Lys Thr Asp Met Val Tyr Leu Glu  
 ACTTTGCAATTAAATGAAATTAACGTATCAAACGGAAATTCTCCACAAATGGAATCGATTGAAATTAGTCAA  
 ▶ Thr Phe Asp Phe Asn Glu Leu Lys Arg Ile Lys Thr Glu Leu Leu Pro Glu Met Glu Met Asp Leu Lys Leu Val Glu  
 TTAAATTGCTATACAGATTGAAAGAAACAAAGAAAAGACCAAAGGGTTATTGGCTAAACTATAATTACGATTGG  
 ▶ Leu Ile Al aTyr Thr Asp Trp Lys Glu Thr Glu Gln Glu Lys Asp Pro Lys Glu Tyr Trp Val Asn Tyr Asn Tyr Asp Trp  
 ATGTTAAACCTGGTGCATGCCGAAAGTGGTTAAATATGCCGATGGTGTGGCCAGGGTGTATATGTTAGTTAA  
 ▶ Met Phe Lys Pro Glu Al aMet Al aGlu Val Val Lys Tyr Al aAsp Glu Val Glu Pro Glu Tyr Trp Tyr Met Leu Val Asn  
 AAAGAAGAAATCCAAACCTGATAATATTGTGTAACCTCGGTGGTAAARGAAACTTGCACAAATATAATGTGGAAGTGCAT  
 ▶ Lys Glu Glu Ser Lys Pro Asp Asn Ile Val Tyr Thr Pro Leu Val Lys Glu Leu Al aGln Tyr Asn Val Glu Val His  
 CCTTACACCGTCCGTAAGATGCACTGGCCGACTTTCACAGACGAAATCAAATGATGCCATTGAA  
 ▶ Pro Tyr Thr Val Arg Lys Asp Al aLeu Pro Glu Phe Phe Thr Asp Val Asn Glu Met Tyr Asp Al aLeu Leu Asn Lys  
 EcoRI Ncol  
 TCAGGGGCAACAGGTGTTACTGATTTCCAGATACTGGGTGGAAATTCTTAAAGGAAATAATCCATGGATGCC  
 ▶ Ser Glu Al a Thr Glu Val Phe Thr Asp Phe Asp Thr Glu Val Glu Phe Leu Lys Glu Ile Lys Ser Met Asp Glu  
 LB1-GR2  
 GGTAAAGCAGGTGTTGCTTACTGCTCTGACTATAATTGCAATAAAAGTAGTAGTAACTGTTACTCTTAA  
 ▶ Glu Lys Al a Glu Val Al a Leu Val Arg Ser Asp Tyr Lys Leu Tyr Asn Lys Asn Ser Ser Asn Ser Thr Leu Lys  
 Ncol  
 AACCTAGGGCAACATCACAGAGCACGTCGCACTGGATGGTGGAAAGCAGGTGTTGCTTTAGTACGTTCTGATTAA  
 ▶ Asn Leu Glu His His Arg Al a Arg Al a Met Asp Glu Val Glu Lys Al a Glu Val Al a Leu Val Arg Ser Asp Tyr Lys  
 LB1-GR1  
 TTATGAAGNTGCAACCGTACTCGTACCCACAAAGAAGGTGCGTACACAGGACGCTACTAGTCGTTCTGACTATAAA  
 ▶ Phe Tyr Glu Asp Al a Asn Glu Thr Arg Asp His Lys Lys Glu Arg His Thr Al a Arg Thr Ser Arg Ser Asp Tyr Lys  
 LB1-GR3 Spel Poly-His XbaI  
 TTCTACGATATAACCGCATGATACTACTAGTGGCCACCATCACCATTAATCTAGATCGATAAGCTTCGA  
 ▶ Phe Tyr Asp Asn Lys Arg Ile Asp Ser Thr Ser Glu His His His His His His His His  
 CCGATGCC

THIS PAGE BLANK (USPTO)

Jia 6

1 2 3 4 5 6 7 8



**THIS PAGE BLANK (USPTO)**

Fig 7

1 2 3 4 5 6 7 8



-200K

-97,4K

-69K



-46K

-30K

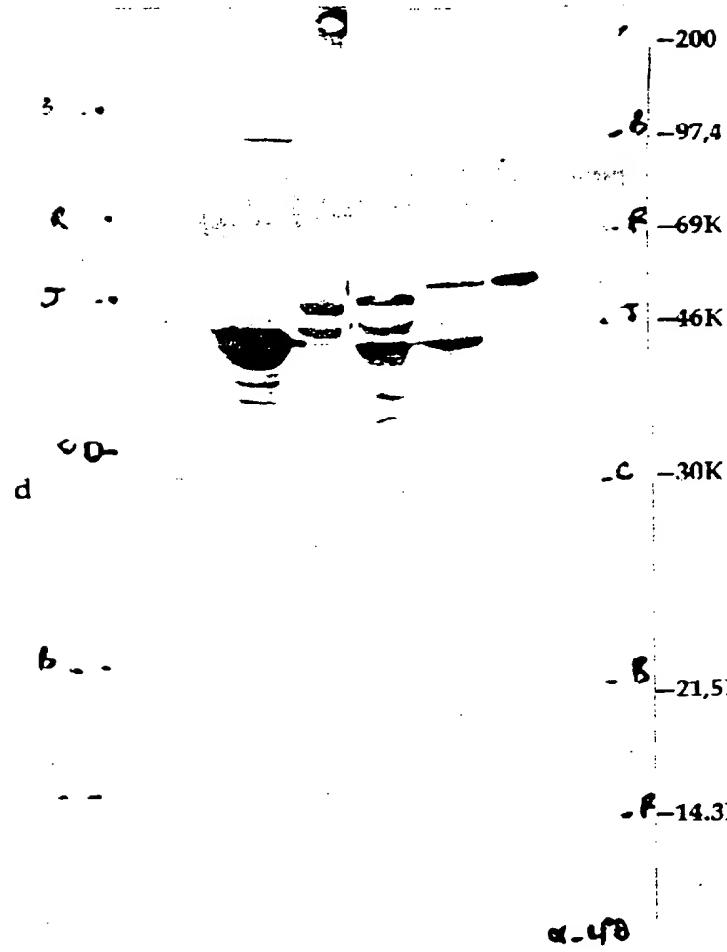
-21,5K

-14,3K

THIS PAGE BLANK (USPTO)

Jig 8

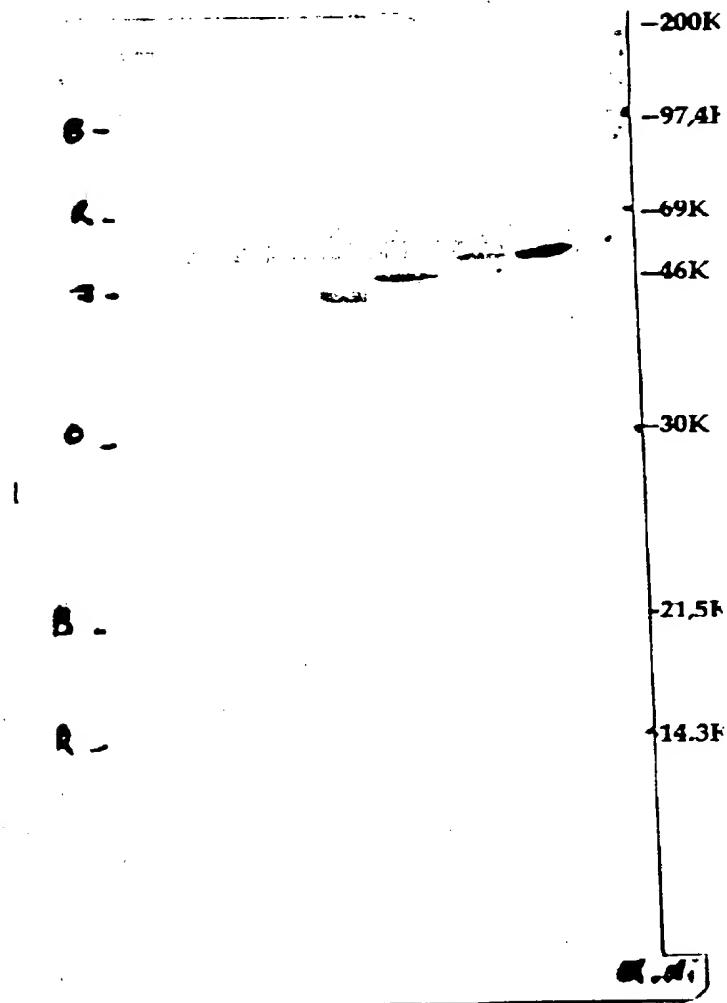
1 2 3 4 5 6 7 8



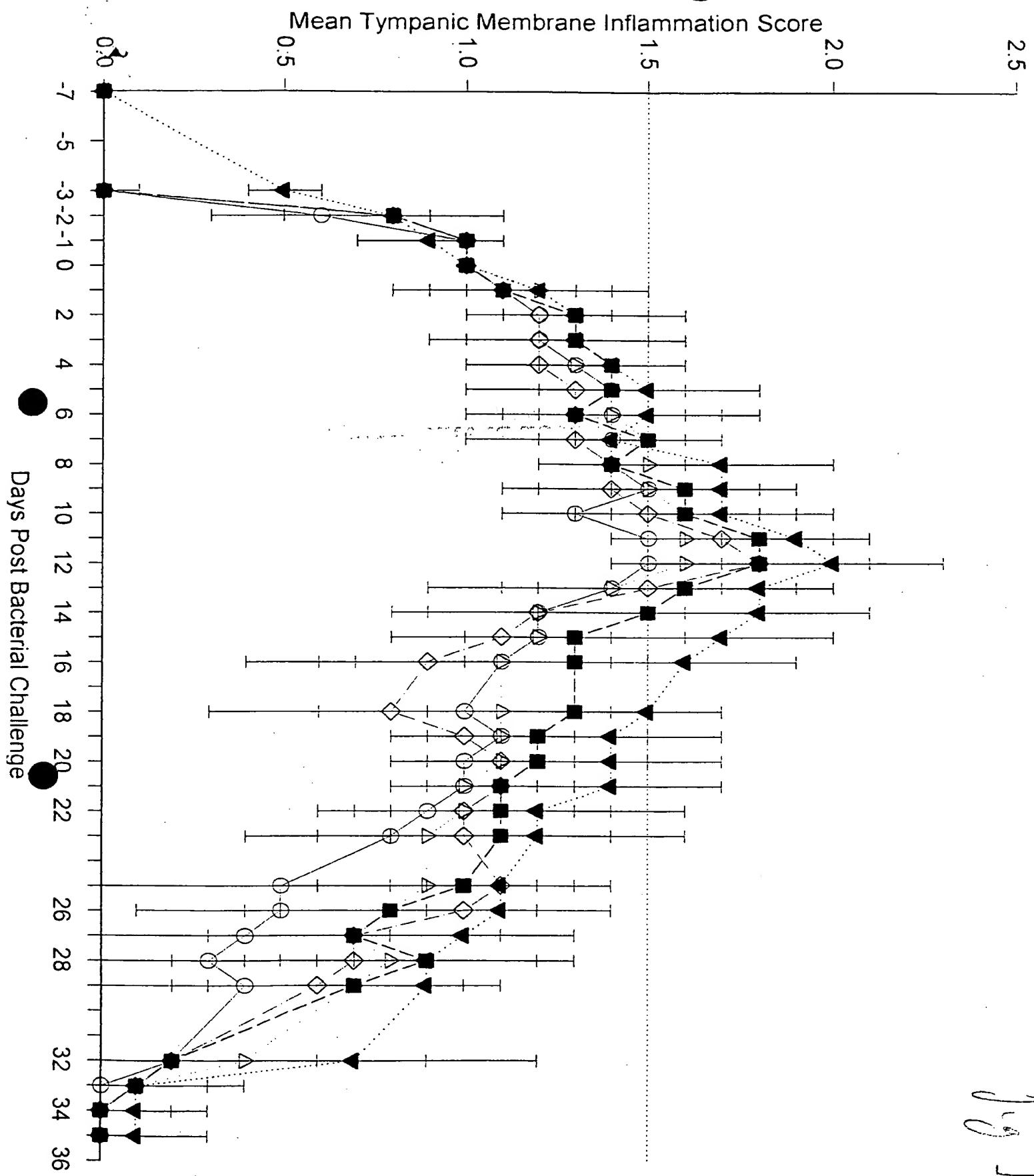
**THIS PAGE BLANK (USPTO)**

Jig 9.

1   2   3   4   5   6   7   8



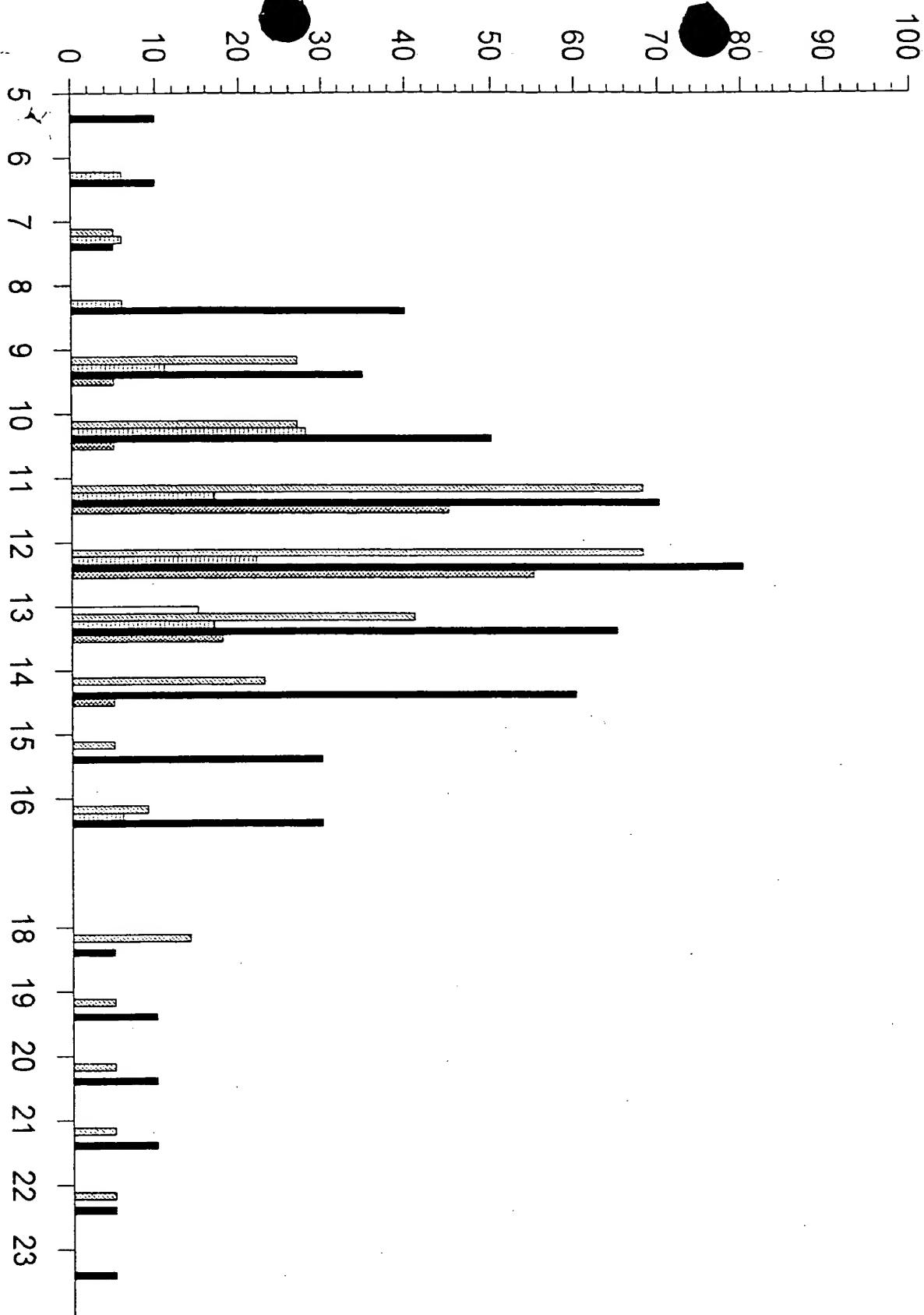
**THIS PAGE BLANK (USPTO)**



10

THIS PAGE BLANK (USPTO)

Percent total ears



11  
11

**THIS PAGE BLANK (USPTO)**